

Syntaktische und phonologische Fehler von DaF-Lernern und ihre Wirkung auf Muttersprachler

INAUGURAL-DISSERTATION

zur

Erlangung des akademischen Grades eines Doktors der Philosophie (Dr. phil.)

im Fachbereich

Germanistik und Kunstwissenschaften

der Philipps-Universität Marburg

vorgelegt von

Atoosa Pazan

aus Teheran

Marburg, Februar 2016

Originaldokument gespeichert auf dem Publikationsserver der
Philipps-Universität Marburg
<http://archiv.ub.uni-marburg.de>



Dieses Werk bzw. Inhalt steht unter einer
Creative Commons
Namensnennung
Weitergabe unter gleichen Bedingungen
4.0 International Lizenz.

Die vollständige Lizenz finden Sie unter:
<https://creativecommons.org/licenses/by-sa/4.0/deed.de>

Vom Fachbereich Germanistik und Kunstwissenschaften
der Philipps-Universität Marburg

als Dissertation angenommen am 17. Februar 2016

Tag der Disputation: 30. August 2016

Betreuer/Erstgutachter: Prof. Dr. Richard Wiese

Zweitgutachterin: Prof. Dr. Ruth Albert

meinen Eltern

Danksagung

Zur Erstellung meiner Dissertation trugen zahlreiche Personen in verschiedener Art und Weise bei. Hiermit möchte ich mich ganz herzlich bei all diesen Personen bedanken.

An erster Stelle gilt mein herzlicher Dank meinem Doktorvater Herrn Prof. Dr. Wiese für seine wissenschaftliche und methodische Unterstützung während der gesamten Bearbeitungsphase. Nur mithilfe seiner zielführenden Ratschläge und konstruktiven Kritik ist es mir gelungen, diese Arbeit in der vorliegenden Form zu fertigen.

Zudem bedanke ich mich ganz herzlich bei meiner Zweitgutachterin Frau Prof. Dr. Albert für ihre fachliche Beratung und Unterstützung.

Besonders möchte ich an dieser Stelle auch meiner Familie für ihren unermüdlichen Zuspruch danken sowie für ihre Geduld und Bereitschaft zu Veränderungen und zum Verzicht im Verlauf meiner Doktorarbeit.

Inhaltsverzeichnis

Abbildungs- und Tabellenverzeichnis.....	I
1. Einleitung.....	1
1.1. Gegenstand und Ziel der Arbeit.....	1
1.2. Überblick über den Inhalt der Kapitel.....	4
2. Theoretische Grundlagen.....	5
2.1. Fehler im Fremdsprachenunterricht.....	5
2.1.1. Auseinandersetzung mit dem Begriff „Fehler“ und seine Rolle im Fremdsprachenunterricht.....	5
2.1.2. Fehlerklassifikation.....	7
2.1.3. Akzent oder Aussprachefehler?.....	11
2.1.4. Wirkung von Sprechfehlern.....	12
2.2. Forschungsstand.....	15
2.2.1. Stand der Forschung zum Thema „Integration aus sprachlicher Sicht“.....	15
2.2.2. Stand der Forschung zum Thema „Die Wirkung des ausländischen Akzents auf die Muttersprachler“.....	16
2.2.3. Stand der Forschung zum Thema „Kontrastive Beobachtung der Wirkung von Aussprache- sowie Grammatikfehlern“.....	19
2.2.4. Stand der Forschung zum Thema „Gewichtung verschiedener Sprachfehler hinsichtlich ihrer Wichtigkeit in der L1-Umgebung“.....	20
3. Vorbereitung und Durchführung der Untersuchung.....	23
3.1. Fragestellung und Hypothesen.....	23
3.2. Auswahl der Fehlertypen.....	27
3.2.1. Auswahlkriterien der Beispiele für Vokalfehler	28
3.2.2. Auswahlkriterien der Beispiele für Konsonantenfehler.....	31
3.3. Design der Untersuchung.....	33
3.3.1. Untersuchungsmaterialien.....	33
3.3.2. Sprecher.....	34
3.3.3. Die Wahl der Stichprobe.....	35
3.3.4. Untersuchungsbeschreibung.....	36

3.3.5	Vorbereitung auf die statistische Auswertung.....	37
4.	Auswertung der Ergebnisse und Überprüfung der Hypothesen.....	42
4.1.	Darstellung der Ergebnisse und Überprüfung der Hypothesen zu den phonologischen Fehlerarten.....	44
4.1.1.	Deskriptive Werte für die Beurteilung der phonologischen Fehler.....	44
4.1.1.1.	Fehlergruppe A: Realisierung der Affrikate [ts]	44
4.1.1.2.	Fehlergruppe B: Auslautverhärtung.....	49
4.1.1.3.	Fehlergruppe C: Vokalquantität.....	51
4.1.1.4.	Fehlergruppe D: Realisierung des Umlauts.....	54
4.1.1.5.	Fehlergruppe E: Konsonantenhäufung in der Wortanlautsilbe.....	57
4.1.1.6.	Fehlergruppe F: Wortakzent.....	60
4.1.2.	Statistische Auswertung der gewonnenen Daten über die Bewertung der phonologischen Fehlerarten.....	63
4.1.2.1.	Varianzanalyse und die Rangordnung der Aussprache-Fehlergruppen....	63
4.1.2.2.	Paarvergleiche bzw. Post-hoc-Analyse der Aussprache-Fehlergruppen...	67
4.1.3.	Überprüfung der Hypothesen zu den phonologischen Fehlern.....	69
4.1.3.1.	Hypothese 1.1.....	69
4.1.3.2.	Hypothese 1.2.....	72
4.1.3.3.	Hypothese 1.3.....	73
4.2.	Darstellung der Ergebnisse und Überprüfung der Hypothesen zu den syntaktischen und morphosyntaktischen Fehlerarten.....	75
4.2.1.	Deskriptive Werte für die Beurteilung der syntaktischen und morphosyntaktischen Fehler.....	75
4.2.1.1.	Fehlergruppe G: Genus.....	75
4.2.1.2.	Fehlergruppe H: Kasus.....	78
4.2.1.3.	Fehlergruppe I: Präposition.....	81
4.2.1.4.	Fehlergruppe J: Partizip II.....	84
4.2.1.5.	Fehlergruppe K: Verbposition im Hauptsatz.....	86
4.2.1.6.	Fehlergruppe L: Verbposition im Nebensatz.....	89
4.2.2.	Statistische Auswertung der gewonnenen Daten über die Bewertung der Syntaktischen und morphosyntaktischen Fehlerarten.....	91
4.2.2.1.	Varianzanalyse und die Rangordnung der Grammatik-Fehlergruppen....	91
4.2.2.2.	Paarvergleiche bzw. Post-hoc-Analyse der Grammatik-Fehlergruppen...	95

4.2.3.	Überprüfung der Hypothesen zu den grammatischen Fehlern.....	96
4.2.3.1.	Hypothese 2.1.....	96
4.2.3.2.	Hypothese 2.2.....	97
4.2.3.3.	Hypothese 2.3.....	99
4.3.	Überprüfung der Hypothesen zum kontrastiven Vergleich der Grammatik- und Aussprachefehler.....	100
4.3.1.	Hypothese 3.....	100
4.3.2.	Hypothese 4.....	103
4.3.2.1.	Hypothese 4.1.....	104
4.3.2.2.	Hypothese 4.2.....	106
4.4.	Überprüfung der Hypothesen zur allgemeinen Wahrnehmung der Sprechfehler.....	109
4.4.1.	Hypothese 5.....	109
4.4.2.	Hypothese 6.....	114
5.	Zusammenfassung und Diskussion der Ergebnisse.....	118
6.	Diskussion und Schlussfolgerung.....	125
	Literaturverzeichnis.....	129

Anhang

Abbildungs- und Tabellenverzeichnis

1	Vergleich der Ergebnisse der Studien von Politzer und Delisle.....	22
2	Die zu untersuchenden grammatischen Abweichungen.....	27
3	Die zu untersuchenden phonologischen Abweichungen.....	27
4	Die Vokalphoneme des Persischen (Hamidi 2006).....	29
5	Die Vokalphoneme des Deutschen (Kürschners 2008: 40)	29
6	Das Konsonantensystem des Deutschen und des Persischen.....	32
7	Angaben zu der Hörergruppe.....	35
8	Statistische Kennwerte für die Fehlergruppe A.....	45
9	Vergleichende Darstellung des Medians und der Streuung für die Gruppe A und die Gruppe mit parallelen korrekten Sätzen.....	46
10	Fehlerangaben der Probanden für die Gruppe A.....	48
11	Absolute und relative Häufigkeit der korrekten Erkennung des Fehlers für die Sätze der Gruppe A.....	48
12	Statistische Kennwerte für die Fehlergruppe B.....	49
13	Vergleichende Darstellung des Medians und der Streuung für die Gruppe B und die Gruppe mit parallelen korrekten Sätzen.....	50
14	Fehlerangaben der Probanden für die Gruppe B.....	51
15	Absolute und relative Häufigkeit der korrekten Erkennung des Fehlers für die Sätze der Gruppe B.....	51
16	Statistische Kennwerte für die Fehlergruppe C.....	52
17	Vergleichende Darstellung des Medians und der Streuung für die Gruppe C und die Gruppe mit parallelen korrekten Sätzen.....	53
18	Fehlerangaben der Probanden für die Gruppe C.....	53
19	Absolute und relative Häufigkeit der korrekten Erkennung des Fehlers für die Sätze der Gruppe C.....	54
20	Statistische Kennwerte für die Fehlergruppe D.....	55
21	Vergleichende Darstellung des Medians und der Streuung für die Gruppe D und die Gruppe mit parallelen korrekten Sätzen.....	56
22	Fehlerangaben der Probanden für die Gruppe D.....	56
23	Absolute und relative Häufigkeit der korrekten Erkennung des Fehlers für	

	die Sätze der Gruppe D.....	57
24	Statistische Kennwerte für die Fehlergruppe E.....	57
25	Vergleichende Darstellung des Medians und der Streuung für die Gruppe E und die Gruppe mit parallelen korrekten Sätzen.....	59
26	Fehlerangaben der Probanden für die Gruppe E.....	59
27	Absolute und relative Häufigkeit der korrekten Erkennung des Fehlers für die Sätze der Gruppe E.....	60
28	Statistische Kennwerte für die Fehlergruppe F.....	61
29	Vergleichende Darstellung des Medians und der Streuung für die Gruppe F und die Gruppe mit parallelen korrekten Sätzen.....	62
30	Fehlerangaben der Probanden für die Gruppe F.....	62
31	Absolute und relative Häufigkeit der korrekten Erkennung des Fehlers für die Sätze der Gruppe F.....	63
32	Friedman-Test: Ranginformation zu den phonologischen Fehlergruppen.....	64
33	Rangordnung der Aussprachefehler hinsichtlich ihrer Wirkung auf die deutschen Muttersprachler.....	65
34	Vergleichende Darstellung des Medians und der Streuung für die phonologischen Fehlerarten.....	65
35	Friedman-Test: Ranginformationen zu den Differenzen der phonologischen Fehlergruppen und parallelen korrekten Sätze.....	66
36	Relative Häufigkeit der korrekten Erkennung der Aussprachefehler und die als ganz sicher angekreuzte Prozentzahl der korrekten Erkennung.....	66
37	Paarweiser Vergleich der Bewertung phonologischer Fehlerarten.....	68
38	Vergleichende Darstellung des Medians und der Streuung für die Mittelwert- Variablen der segmentalen und suprasegmentalen Fehlerarten.....	71
39	Vergleichende Darstellung des Medians und der Streuung für die Mittelwert- Variablen der vokalischen und konsonantischen Fehlergruppen.....	73
40	Statistische Kennwerte für die Fehlergruppe G.....	76
41	Vergleichende Darstellung des Medians und der Streuung für die Gruppe G und die Gruppe mit parallelen korrekten Sätzen.....	77
42	Fehlerangaben der Probanden für die Gruppe G.....	77

43	Absolute und relative Häufigkeit der korrekten Erkennung des Fehlers für die Sätze der Gruppe G.....	78
44	Statistische Kennwerte für die Fehlergruppe H.....	79
45	Vergleichende Darstellung des Medians und der Streuung für die Gruppe H und die Gruppe mit parallelen korrekten Sätzen.....	80
46	Fehlerangaben der Probanden für die Gruppe H.....	80
47	Absolute und relative Häufigkeit der korrekten Erkennung des Fehlers für die Sätze der Gruppe H.....	81
48	Statistische Kennwerte für die Fehlergruppe I.....	82
49	Vergleichende Darstellung des Medians und der Streuung für die Gruppe I und die Gruppe mit parallelen korrekten Sätzen.....	83
50	Fehlerangaben der Probanden für die Gruppe I.....	83
51	Absolute und relative Häufigkeit der korrekten Erkennung des Fehlers für die Sätze der Gruppe I.....	84
52	Statistische Kennwerte für die Fehlergruppe J.....	85
53	Vergleichende Darstellung des Medians und der Streuung für die Gruppe J und die Gruppe mit parallelen korrekten Sätzen.....	85
54	Fehlerangaben der Probanden für die Gruppe J.....	86
55	Absolute und relative Häufigkeit der korrekten Erkennung des Fehlers für die Sätze der Gruppe J.....	86
56	Statistische Kennwerte für die Fehlergruppe K.....	87
57	Vergleichende Darstellung des Medians und der Streuung für die Gruppe K und die Gruppe mit parallelen korrekten Sätzen.....	88
58	Fehlerangaben der Probanden für die Gruppe K.....	88
59	Absolute und relative Häufigkeit der korrekten Erkennung des Fehlers für die Sätze der Gruppe K.....	89
60	Statistische Kennwerte für die Fehlergruppe L.....	89
61	Vergleichende Darstellung des Medians und der Streuung für die Gruppe L und die Gruppe mit parallelen korrekten Sätzen.....	90
62	Fehlerangaben der Probanden für die Gruppe L.....	91
63	Absolute und relative Häufigkeit der korrekten Erkennung des Fehlers für	

	die Sätze der Gruppe L.....	91
64	Friedman-Test: Ranginformation zu den Grammatik-Fehlergruppen.....	92
65	Rangordnung der Grammatikfehler hinsichtlich ihrer Wirkung auf die deutschen Muttersprachler.....	92
66	Vergleichende Darstellung des Medians und der Streuung für die grammatischen Fehlerarten.....	93
67	Friedman-Test: Ranginformation zu den Differenzen der grammatikalischen Fehlergruppen und parallelen korrekten Sätze.....	94
68	Relative Häufigkeit der korrekten Erkennung der Grammatikfehler und die als ganz sicher angekreuzte Prozentzahl der korrekten Erkennung.....	94
69	Paarweiser Vergleich der Bewertung grammatischer Fehlerarten.....	95
70	Vergleichende Darstellung des Medians und der Streuung für die phonologischen und grammatischen Fehlerarten.....	100
71	Vergleichende Darstellung des Medians und der Streuung für die Mittelwert- Variablen der phonologischen und grammatischen Fehlerarten.....	101
72	Vergleichende Darstellung des Medians und der Streuung für die Differenz- Variablen der phonologischen und grammatischen Fehlerarten.....	102
73	Häufigkeitsverteilung der Fehlerangaben für die Gruppe der Aussprachefehlerarten.....	105
74	Häufigkeitsverteilung der Fehlerangaben für die Gruppe der Grammatikfehlerarten.....	105
75	Häufigkeitsverteilung der Sicherheitsangaben bei der Erkennung der Aussprachefehler.....	107
76	Häufigkeitsverteilung der Sicherheitsangaben bei der Erkennung der Grammatikfehler.....	107
77	Mittelwertvergleich der zwei Ausprägungen der Fehlererkennung.....	111
78	Korrelation zwischen der Bewertung und Fehlererkennung in den Grammatik- sowie Aussprache-Fehlergruppen.....	112
79	Korrelation zwischen der Bewertung und Fehlererkennung in den Untergruppen der Grammatik- und Aussprache-Fehlergruppen.....	113
80	Häufigkeitsverteilung der Fehlerangaben für die korrekten Sätze.....	116

81	Rangliste der Aussprache- und Grammatik-Fehlerarten.....	118
82	Zusammenstellung der segmentalen und suprasegmentalen Fehlergruppen.....	119

1. Einleitung

1.1. Gegenstand und Ziel der Arbeit

„Zuwanderungsgesetz – Integrationskurse – Integrationsgipfel: So viel ‚Integration‘ wie derzeit gab es wohl noch nie, jedenfalls nicht auf Seiten der offiziellen Politik.“ Die vor acht Jahren geäußerte sorgenvolle Aussage von Barkowski (2008:14) über das Problem der Integration von Migranten in Deutschland passt sehr gut zu der heutigen Situation hinsichtlich der Flüchtlingskrise. Angesichts der stark ansteigenden Anzahl der Flüchtlinge gehört das Thema „Die Gründe für das Scheitern der Integration von Migranten“ zu einem der aktuellsten Debattier-Themen in der Bundesrepublik Deutschland. Neben vielen Faktoren, die den Integrationsprozess beeinflussen, spielen die Kenntnisse bzw. die Beherrschung der deutschen Sprache eine zentrale Rolle. Die Sprache gilt insofern als eine der grundlegendsten Voraussetzungen für Integration, als sie als ein kommunikatives Instrument die Erkennung und die Aufnahme der andersartigen sprachlichen und kulturellen Gegebenheiten ermöglicht.

Die Beantwortung der Frage, wie schnell bzw. effektiv und mit welchen Mitteln ein Migrant eine ausreichende Sprachkompetenz als grundlegende Fähigkeit für Integration erwirbt¹, steht im Mittelpunkt des wissenschaftlichen Interesses der Fremdsprachendidaktik bzw. Sprachlehrforschung. Um dieser Frage nachzugehen und die Qualität der Deutschkurse zu verbessern, führen Fremdsprachendidaktiker viele empirische Forschungen über die Theorie und Praxis des Fremdsprachenunterrichts durch. Eines ihrer wichtigsten Forschungsgebiete ist die Beobachtung und Analyse des Fremdsprachenerwerbsprozesses. Um dies zu erforschen, beschäftigen sie sich neben der Untersuchung psychologischer und sozialer Aspekte des Spracherwerbs mit der Erfassung und Dokumentation des Lernstandes bzw. Lernfortschritts der Lernenden.

Bei der Abschätzung des Stands der Lerner Sprache spielen die sprachlichen Fehlleistungen eine wichtige Rolle. Während in der Vergangenheit einige Linguisten wie Kielhöfer (1975:21) Fehler als „Defekte“ bezeichneten, deren Bekämpfung und Ausmerzung die Hauptaufgabe des Fremdsprachenunterrichts sei, betrachten heute die

¹ Obwohl Erwerben in der Literatur oft als Aneignung der Sprache in einer natürlichen Kommunikation und Lernen als Aneignung der Sprache durch Unterricht definiert wird, gehen wir in der vorliegenden Arbeit nicht auf diese Differenzierung ein, sondern verwenden beide Begriffe als Synonyme, wobei die Bedeutung von Lernen durch Unterricht im Mittelpunkt unserer Arbeit steht (vgl. Barkowski & Krumm 2010:191).

Fremdsprachendidaktiker Fehler eher als ein Signal, das uns wichtige Hinweise auf den Stand des Lernprozesses liefert. Diese positive Einstellung zum Fehler ist vor allem auf Corder (1978:25) zurückzuführen. Seiner Ansicht nach können Fehler das aktuelle System der Lernaltersprache reflektieren und bieten den Lehrenden, Lernenden und Forschern brauchbare Informationen. Die Lehrer können durch Beobachtung der Fehler erkennen, inwieweit die Lerner ihre Sprachkenntnisse verbessert haben und welche Themen mit welchen Methoden noch einzusetzen sind. Die Lerner verwenden Fehler als Lernstrategie und können so ihre Hypothesen über die Zielsprache überprüfen. Die Beobachtung der sprachlichen Fehlleistungen ermöglicht den Forschern außerdem, den Lernprozess und die Lernstrategien zu analysieren.

Eine systematische Analyse der Fehler vollzieht sich durch die Beschreibung und Klassifikation der Fehler aus verschiedenen Perspektiven. Spillner (1991:XXVI) fasst die Schritte zur Fehleranalyse folgendermaßen zusammen:

- Fehleridentifizierung (Lokalisierung, Normdiskussion, Ermittlung der Kommunikationsintention)
- Fehlerdeskription (Beschreibung und Klassifikation nach linguistischen Kriterien)
- Hypothesen zur Fehlerursache (psychologische oder psycholinguistische Interpretation der wahrscheinlichen Fehlerquellen)
- Ermittlung der kommunikativen Wirkung (Interpretation der Konsequenzen von Fehlern im Kommunikationsprozess)
- Fehlerdidaktik (Fehlerevaluation, Fehlertherapie, Fehlerprophylaxe)

Während die produzierte Form der sprachlichen Abweichungen, nämlich die Identifizierung, Beschreibung und Klassifikation nach linguistischen Ebenen, sowie die Ursachen und der didaktische Umgang mit Fehlern relativ gut erforscht und analysiert sind, gibt es wenige Untersuchungen, die sich mit der Klassifikation der Fehler nach ihrer Wirkung auf die Muttersprachler bzw. aus rezeptiver Sicht beschäftigen. Angesichts der Tatsache, dass nicht alle sprachlichen Abweichungen gleichermaßen negativ wirken und letztendlich der muttersprachliche Rezipient darüber entscheidet, ob eine Äußerung Sprachfehler beinhaltet, kann die Analyse der Wirkung sprachlicher Fehler einen wichtigen Beitrag zur Entwicklung des Fremdsprachenunterrichts leisten.

Basierend auf diesem Hintergrund bildet die Klassifikation und Gewichtung der gesprochenen Fehler von L2-Sprechern nach ihrer störenden Wirkung auf deutsche Muttersprachler den Gegenstand der vorliegenden Untersuchung. Diese Arbeit setzt sich an allererster Stelle das Ziel, durch die kontrastive Beobachtung der Wirkung verschiedener Fehlerarten auf Deutsche folgende Fragen zu beantworten: Die Beherrschung welcher Kenntnisse spielt für die sprachliche Akzeptanz in der L1-Gesellschaft eine wichtigere Rolle, die der Grammatik oder der Aussprache? Ist die Toleranz der L1-Sprecher gegenüber verschiedenen Fehlerarten unterschiedlich? Wenn ja, welche Fehlerarten fallen vornehmlich auf und werden negativer sanktioniert? Inwieweit kann ein ausländischer Akzent zu einer falschen Beurteilung der Sprachkenntnisse führen?

Die Beantwortung dieser Fragen, die durch eine empirische Untersuchung über die Reaktion deutscher Muttersprachler auf verschiedene grammatische und phonologische Sprechfehler eines Migranten herausgearbeitet wird, soll Aufschluss darüber geben, wie gesprochene Fehlerarten hinsichtlich ihres Störungsgrads für deutsche Rezipienten zu gewichten sind. Diese Klassifikation sollte vor allem DaF-Lehrkräfte und -Lernende dazu befähigen, für die Therapie und Behebung der Fehlerarten Prioritätskriterien zu entwickeln.

1.2. Überblick über den Inhalt der Kapitel

Kapitel 2 der vorliegenden Arbeit setzt sich mit den theoretischen Grundlagen auseinander, die für unsere empirische Untersuchung von großer Bedeutung sind. Dabei gehen wir auf wichtige Aspekte des Sprachfehlers wie dessen Definition und Klassifikation sowie seine Wirkung auf die Rezipienten ein und versuchen die Frage zu beantworten, ob und inwiefern ein fremder Akzent als Aussprachefehler zu realisieren ist. Anschließend stellen wir den aktuellen Forschungsstand dar, um die Wichtigkeit unserer Untersuchung in Bezug auf die noch offen gebliebenen Fragen zu verdeutlichen.

Kapitel 3 führt mit der Vorstellung der aufgestellten Fragestellungen und Hypothesen in den empirischen Teil ein. In ihm wird zunächst auf die Kriterien für die Auswahl der Fehlerarten und zugehörige Beispiele eingegangen. Dabei stützen wir uns für die phonologischen Fehlerarten auf die kontrastive Beobachtung der Ausspracheeigenheiten der Sprachen Persisch und Deutsch und für die grammatischen Fehlerbeispiele auf die von DaF-Lernern oft begangenen Fehlerarten. Im Weiteren stellen wir die Untersuchungsaspekte wie die Testmaterialien, den Sprecher und die Hörer vor und beschreiben die Durchführung der Untersuchung sowie die Art und Weise, wie wir unsere gewonnenen Daten für die weitere statistische Auswertung vorbereitet haben.

Im Kapitel 4, das den breitesten Raum unserer Arbeit einnimmt, geht es um die Auswertung der Ergebnisse und die Überprüfung der aufgestellten Fragestellungen und Annahmen. Für die Prüfung der Hypothesen zu den phonologischen und grammatischen Fehlerarten beschreiben wir zunächst für die einzelnen Fehlerarten anhand geeigneter Tabellen und Abbildungen die gewonnenen deskriptiven Daten und vergleichen diese Maßzahlen mit der äquivalenten Bewertung der korrekten Sätze. Mittels dieser deskriptiven Vorauswertung erkennen wir mögliche Unterschiede und Zusammenhänge zwischen den Fehlerarten, die wir anschließend für die Überprüfung einzelner Hypothesen anhand passender Testverfahren auf Signifikanz überprüfen können.

Kapitel 5 fasst die Ergebnisse der durchgeführten empirischen Untersuchung zu den einzelnen Hypothesen zusammen und die sich im Kapitel 6 anschließende Diskussion umfasst eine kritische Betrachtung der Untersuchungsmethode und gibt einen Überblick über die erzielten, für den DaF-Unterricht bedeutenden Ergebnisse sowie die Forschungsdefizite in Bezug auf die noch offen gebliebenen Fragen.

2. Theoretische Grundlagen

2.1. Fehler im Fremdsprachenunterricht

Da die Analyse der Wirkung von sprachlichen Fehlern im Mittelpunkt der vorliegenden Arbeit steht, ist es an erster Stelle wichtig zu bestimmen, wie der Begriff „Sprachfehler“ definiert und nach welchen Kriterien er klassifiziert wird. Im Hinblick auf die Zielsetzung der folgenden Untersuchung gehen wir außerdem auf die Fragen ein, welche Wirkung die sprachlichen Fehler auf die L1-Sprecher ausüben und welche Rolle diese Wirkung beim Erfolg der Kommunikation spielt.

2.1.1. Auseinandersetzung mit dem Begriff „Fehler“ und seine Rolle im Fremdsprachenunterricht

Mein Unterrichtskonzept für die erste Sitzung des Deutschkurses für Mittelstufe beinhaltete unter anderen formalen Einführungen immer eine Aufgabe zur Fehlerkorrektur. Die Übung bestand darin, dass die Teilnehmer die sprachlichen Fehler in Sätzen, die von anderen Lernern gleicher Stufe produziert worden waren, identifizieren und korrigieren. Das positive Feedback der Teilnehmer und ihre Begeisterung über diese Übung in jedem Kurs bestätigte die motivierende Wirkung der Fähigkeit zur Fehlererkennung. Viele Teilnehmer, die über die kommunikative Kompetenz, sich problemlos mit Deutschen zu verständigen, verfügten, gaben nach der Durchführung dieser Übung an, dass ihr Hauptziel im Deutschunterricht fehlerfreies Sprechen und Schreiben sei. Diese Äußerung widerspricht der Einsicht einiger Linguisten wie Huber (1974) und Andresen (1976, zit. nach Kolde 1980:183), die die Verständigung als einziges Ziel des Sprachhandels betrachten.

Wie wir aber aus der Praxis des Fremdsprachenerwerbs der Erwachsenen wissen, ist das Erreichen eines sowohl phonologischen als auch grammatischen L1-Niveaus für L2-Lerner fast unmöglich. Die Fremdsprachenlehrer setzen vor allem darauf, die Anzahl der Fehler in den sprachlichen Äußerungen der Lernenden zu minimieren. Nun stellt sich die Frage, nach welchen Kriterien wir eine sprachliche Aussage als falsch oder richtig bezeichnen. Die Unterscheidung zwischen korrekt und inkorrekt, die eine schwierige Aufgabe für die Lehrkräfte im Fremdsprachenunterricht ist, hängt von vielen Faktoren wie dem Sprachniveau der Lernenden, ihrer individuellen Lernziele und Lebenssituation sowie ihrem Alter ab. Eine

Aussage wie *Ich habe vier Praktikum gemacht*. wird z. B. für einen Lerner mit dem Niveau A2 als weniger fehlerhaft bezeichnet als für einen Lerner in der C1-Stufe.

Diese relative Bedeutung des Fehlerbegriffs gibt es allerdings nur in der Unterrichtssituation. In einer Kommunikationssituation außerhalb des Unterrichts bleibt die individuelle Lernsituation der Lernenden meist außer Berücksichtigung. Die Beantwortung der Frage „Was ist ein Sprachfehler in einer realen Kommunikationssituation?“², die auf den ersten Blick sehr trivial erscheint, bereitet Linguisten einige Schwierigkeiten. Bußmann (2008:4) definiert den Begriff „sprachliche Abweichung“ als eine „Eigenschaft von Ausdrücken einer natürlichen Sprache, die entweder nicht im Einklang stehen mit stillschweigend oder explizit vereinbarten Übereinkünften (→Sprachnorm) oder mit linguistischen Beschreibungen (→Regeln).“

Die Abgrenzung der Termini „Sprachnorm“ und „Regeln bzw. Sprachsystem“ als Maßstab für die Bestimmung der Korrektheit einer sprachlichen Äußerung, die auf Coseriu (1975) zurückgeht, wird jedoch von einigen Linguisten wie Kielhöfer (1975:55) als problematisch betrachtet. Er hält die Unterscheidung zwischen System und Norm als nicht eindeutig durchführbar und meint:

Das System stellt die sprachlichen Möglichkeiten dar, den virtuellen Bereich, von denen ein Teil auf der Normebene realisiert wird. Was aber wirklich realisierbar ist, kann letzten Endes nur durch den Sprachgebrauch der native speaker entschieden werden.

Insofern wird der Begriff „Sprachnorm“ als durch die Gesellschaft akzeptierte Realisierung des Sprachsystems definiert (vgl. Szarska-Wieruszewska 2005:31). Für die Bestimmung der Korrektheit oder Fehlerhaftigkeit der sprachlichen Äußerung im Fremdsprachenunterricht sind deshalb System und Norm als komplementäre Kriterien zu betrachten.

Kleppin (1998:20) zieht neben dem Kriterium „Korrektheit“, das sich als Anpassung der fremdsprachlichen Äußerungen an das System und die Norm einer Zielsprache verstehen

² Die Abweichung von einer sprachlichen Norm kann sich auch in den Äußerungen der L1-Sprecher in verschiedenen Formen wie Versprechern manifestieren. Auf diese Fehlerdimension gehen wir an dieser Stelle aus Umfangsgründen, aber auch im Hinblick auf die Zielstellung unserer Untersuchung nicht ein. Der Terminus „Sprachfehler“ bezieht sich also in der vorliegenden Arbeit nur auf die fehlerhaften Sprachäußerungen der L2-Sprecher.

lässt, für die Identifizierung der Fehler andere Kriterien wie „Verständlichkeit“ und „Situationsangemessenheit“ heran. Hier stellt sich die Frage, wer bestimmt, ob eine sprachliche Äußerung passend zu dem System und der Norm der Zielsprache, verständlich und in einer bestimmten Kommunikationssituation angemessen ist, und wie das ermittelt wird. Wie wir wissen, sind zwei der wichtigsten Grundkomponenten des Kommunikationsmodells der Sender und der Empfänger bzw. Sprecher und Hörer (vgl. Bußmann 2008:347). Die Äußerung des Sprechers wird vor allem vom Rezipienten verstanden und hinsichtlich ihrer Fehlerhaftigkeit beurteilt.

Da die Beurteilung der sprachlichen Fehler, die keine Verständigungsprobleme verursachen, aus der Perspektive der muttersprachlichen Hörer im Mittelpunkt der vorliegenden Arbeit steht, definieren wir den Begriff „Fehler“ als eine Lerneräußerung, die von der Mehrheit der muttersprachlichen Rezipienten als abweichend von der Standardsprache diagnostiziert wird, und nehmen auf die Definition von Kolde (1980:172) Bezug, die besagt:

Ein Merkmal einer Äußerung eines Sprechers S gilt einem Rezipienten R dieser Äußerung als *sprachlicher Fehler*, also als hör- oder sichtbare Folge einer sprachlichen Fehlhandlung, wenn R dieses Merkmal darauf zurückführt, daß (dass) S eine für ihn „obligate“ sprachliche Norm oder Regel nicht befolgt hat, sei es, weil er sie nicht beherrscht (=anzuwenden gelernt hat), sei es, obwohl er sie beherrscht.

2.1.2. Fehlerklassifikation

Neben der Identifizierung der Fehler ist die Bestimmung der Art ihrer Fehlerhaftigkeit eine wichtige Aufgabe der Fehleranalyse. Die Klassifikation der Fehler nach dem Grad ihrer Fehlerhaftigkeit und Kommunikationsstörung, nach ihren Ursachen sowie ihrer Zugehörigkeit zu einer linguistischen Ebene hilft vor allem den Lehrkräften, bei der Bewertung und Therapie der Fehler eine Korrekturstrategie, die bei der entsprechenden Fehlerart besser passt, anzuwenden. Für die Aufteilung der Fehler in verschiedene Kategorien wird in der Fachliteratur eine ganze Reihe verschiedener Kriterien genannt. Da die Beschreibung aller Klassifikationsmöglichkeiten den Umfang dieser Arbeit sprengen würde, beschränken wir uns im Folgenden auf die Vorstellung von Aufteilungskriterien, die für die Auswahl der Fehler in unserer Untersuchung von Relevanz sind und von den meisten Linguisten besondere Beachtung geschenkt bekommen.

Grad der Fehlerhaftigkeit

Die Aufteilung der Fehler nach dem Grad ihrer Fehlerhaftigkeit geht vor allem auf die von Corder (1974:25) eingeführte Unterscheidung zwischen Kompetenz- und Performanzfehler zurück. Die Kompetenzfehler, die er auch als *error* oder systematische Fehler bezeichnet, sind solche, die aufgrund der sprachlichen Unwissenheit der Lernenden zustande kommen. Bei diesen Fehlern hat der Lerner entweder eine sprachliche Regel noch nicht gelernt oder falsch verstanden und kann sie deshalb nicht selbstständig korrigieren. Die Performanzfehler, die auch *mistake* genannt werden, beziehen sich dagegen auf Fehler, die nicht aufgrund des Mangels an Sprachkompetenz, sondern infolge anderer Faktoren wie Konzentrationsmangel eintreten. Diese Fehler, die selbst in zwei Gruppen, nämlich Fehler, die aufgrund der mangelnden Übung passieren, und Versprecher, aufgeteilt werden, können von den Lernenden selbst erkannt und korrigiert werden (vgl. Kleppin 1998). Die Einteilung der Fehler in diese zwei Gruppen ist jedoch nur in der Theorie vorstellbar. In der Praxis ist die Entscheidung darüber, ob eine fehlerhafte Äußerung des L2-Sprechers auf seine Unkenntnisse der Sprachstruktur oder auf oberflächliche Ursachen, z. B. Konzentrationsmangel, zurückzuführen ist, schwer zu realisieren (vgl. Raabe 1980:81).

Grad der Kommunikationsstörung

Ein weiteres Aufteilungskriterium für Fehler bezieht sich auf den Grad ihrer Kommunikationsstörung. Kleppin (1998:42) bezeichnet Fehler, die den Verständigungsprozess erschweren, als kommunikationsbehindernd und stellt sie den nicht kommunikationsbehindernden Fehlern gegenüber, die keine Verständigungsprobleme verursachen. Burt (1975) unterscheidet in dieser Hinsicht zwischen globalen und lokalen Fehlern. Als global bezeichnet sie Fehler, die das Verständnis einer Äußerung beeinträchtigen und als lokal betrachtet sie Abweichungen, die keinen Einfluss auf den Verständigungsprozess ausüben. Sie argumentiert weiterhin dafür, dass die globalen Fehler in der gesprochenen Sprache schwerwiegender auffallen als die lokalen. Wenn wir aber die Tatsache, dass einige schwerwiegende Grammatikfehler für eine erfolgreiche Verständigung in der Kommunikation in keiner Weise von Bedeutung sind, berücksichtigen, können wir davon ausgehen, dass ein gelungener oder gescheiterter Verständigungsprozess nicht unbedingt als Maßstab für die Bestimmung der Schwere der Fehler angesehen werden darf (vgl. Kleppin 1998:42).

Fehlerursachen

Um die Ursachen der Sprachfehler von L2-Sprechern analysieren zu können, verwenden Linguisten unterschiedliche Herangehensweisen. Ein großer Anteil der Arbeiten, die bis jetzt durchgeführt wurden, beschäftigte sich mit der Analyse der Unterschiede zwischen der Mutter- und der Zielsprache und begründete so die Fehlerursache mit der Übertragung der muttersprachlichen Sprachgewohnheiten bzw. -strukturen, die in der Zielsprache keine Äquivalenz finden, auf die fremde Sprache. Der kontrastive Vergleich zweier Sprachen kann allerdings nur eine begrenzte Anzahl von Fehlern diagnostizieren. Angesichts der Tatsache, dass der L2-Lerner im Spracherwerbsprozess als „kreativ“ und „kognitiv“ tätig gesehen wird, können wir seine Fehler nicht nur auf seine Muttersprache, sondern auch auf andere und manchmal sogar auf mehrere Ursachen zurückführen (vgl. Kleppin 1998:39). Im Folgenden präsentieren wir die wichtigsten Ursachen, die in der Fachliteratur genannt werden (vgl. Kleppin 1998, Katsikas 1995). Bei dieser Darstellung teilen wir die Fehlerursachen in drei Gruppen auf, nämlich Fehler, die aufgrund des Einflusses der Zielsprache, des individuellen Sprachgebrauchs der Lernenden und des Fremdsprachenunterrichts zustande kommen. Neben diesen Ursachen können auch persönliche bzw. psychologische Faktoren wie Angst, Unsicherheit oder Mangel an Aufmerksamkeit sowie soziokulturelle Faktoren zur Entstehung von Fehlern beitragen, auf deren Ausführung wir an dieser Stelle aus Umfangsgründen verzichten.

Die wichtigsten Fehler, die infolge des Einflusses der Zielsprache vorkommen, sind Übergeneralisierung, Vereinfachung und Regularisierung. Die Übergeneralisierung passiert, wenn eine bereits gelernte Regel der Fremdsprache sich fälschlicherweise auf andere unpassende Strukturen ausweitet, z. B. die Verwendung der falschen Partizip-II-Formen wie *untergesucht* statt *untersucht*. Vereinfachung bzw. Simplifizierung kommt zustande, wenn einige schwierige Elemente der Sprache reduziert und sogar vermieden werden, z. B. das Weglassen der Artikel oder Adjektivendungen. Als Regularisierung bezeichnet man die Änderung einer unregelmäßigen Regel zu einer regelmäßigen, z. B. die Verwendung des Verbs *gehte* anstelle von *ging*.

Fehler, die auf den individuellen Sprachgebrauch der Lerner zurückzuführen sind, passieren unter dem Einfluss von Kommunikations- und Lernstrategien. Unter Kommunikationsstrategien sind individuelle Methoden zu verstehen, die ein Lerner, der die Zielsprache noch nicht beherrscht, benutzt, um erfolgreich mit anderen zu kommunizieren.

Kleppin (1998:34) präsentiert in dieser Hinsicht eine Reihe von Beispielen wie die Umschreibung eines Begriffs, die Verwendung muttersprachlicher Ausdrücke oder die Erfindung neuer Komposita, die zur Entstehung und sogar Fossilisierung der Fehler führen können. Die persönlichen Lernstrategien, die als Untergruppe der Kommunikationsstrategien betrachtet werden, beziehen sich auf kreatives Ausprobieren neuer sprachlicher Strukturen, die wiederum zur Bildung fehlerhafter Strukturen beitragen können.

Fehler können schließlich auch durch Elemente des Fremdsprachenunterrichts hervorgerufen werden, z. B. Fehler, die durch unklar formulierte Drillübungen verursacht werden, wobei das Beispiel für die Lösung der Aufgabe bzw. der Mustersatz nicht auf alle Sätze übertragbar ist und so zu einer falschen Formulierung führt.

Klassifikation nach linguistischen Ebenen

Ein wichtiger Klassifikationsansatz, der für unsere Arbeit besonders relevant ist, ist die Kategorisierung der Fehler aus linguistischer Sicht. Die fehlerhaften Äußerungen können auf verschiedene linguistische Ebenen, nämlich Phonologie, Morphologie, Syntax, Semantik oder Pragmatik zurückgeführt werden. Kleppin (1998:42-43) klassifiziert Fehler in diesem Zusammenhang nach Sprachebenen und präsentiert folgende Fehlertypen: „phonetisch/phonologische, morphosyntaktische, lexikosemantische, pragmatische und inhaltliche Fehler“. Da die linguistischen Ebenen sich manchmal überlappen, tauchen allerdings bei dieser Aufteilung oft Probleme auf. Es existiert eine Reihe von Fehlern, die nicht nur einer, sondern mehreren linguistischen Kategorien zugeordnet werden können. Die phonologische Erscheinung „Umlaut“, die in der Literatur zumeist als ein rein phonologisches Phänomen betrachtet wird, kann z. B. auch als eine morphologische Erscheinung angesehen werden (vgl. Wiese 1987). Auch innerhalb einer linguistischen Ebene kommt es manchmal vor, dass man zwischen einigen Fehlertypen nicht einfach unterscheiden kann, z. B. zwischen Genus- und Kasusfehlern. Trotz dieser Schwierigkeiten bei der Zuordnung der Fehlerart bietet die linguistische Kategorisierung eine relativ klare Richtlinie für die Beschreibung und Definition des Fehlers an.

In der vorliegenden Arbeit orientieren wir uns an einer linguistischen Fehlerklassifikation und beobachten die Wirkung von verschiedenen Fehlertypen, die den zwei sprachlichen Ebenen, nämlich der Aussprache bzw. Phonologie und der Grammatik bzw. Syntax und Morphosyntax, zugeordnet sind. In beiden Kategorien verstehen wir Fehler als

Abweichungen, die keine Verständigungsprobleme verursachen und infolge des Einflusses der Mutter- und Zielsprache zustande kommen. Die Bewertung der Fehler setzt voraus, dass die Probanden erkennen, dass der Sprecher eine sprachliche Norm oder Regel nicht berücksichtigt hat. Angesichts der Tatsache, dass in unserer Untersuchung der Sprecher mit einem leichten fremden Akzent spricht, scheint die Erkennung der Aussprachefehler im Gegensatz zu den Grammatikfehlern nicht einfach zu sein. Da die Unterscheidung zwischen dem Aussprachefehler und einem fremdartigen Akzent für unsere Untersuchung von besonderer Bedeutung ist, gehen wir im Folgenden näher auf diese Differenzierung ein.

2.1.3. Akzent oder Aussprachefehler?

Der fremde Akzent wird in der Fremdsprachendidaktik als „individuelle Ausspracheeigenheit“ bezeichnet, die durch Übertragung der „Artikulationsgewohnheiten der Erstsprache“ auf die Zielsprache zustande kommt (vgl. Bußmann 2008:23). Unter Aussprachefehlern bzw. phonetischen Fehlern sind Erscheinungen zu verstehen, die eine Abweichung von der Standardaussprache aufweisen (vgl. Rausch 1988:46). Die Standardaussprache bezieht sich auf Ausspracheformen, die in Standardwerken wie dem „Wörterbuch der deutschen Aussprache“ oder dem „Duden-Aussprachewörterbuch“ beschrieben werden und die Aussprache der Rundfunk- und Fernsehsprecher widerspiegeln (vgl. Duden-Aussprachewörterbuch 2005:34).

Obwohl Akzent und Aussprachefehler beide als eine Abweichung von der Normaussprache zu verstehen sind, können sie nicht als synonyme Begriffe verwendet werden. Dieser Unterschied kann vor allem in den Bewertungen der L1-Sprecher über die Sprachkompetenz der L2-Lerner beobachtet werden. Folgende Beurteilungen hören wir oft von Deutschen:

- A. Er spricht sehr gut Deutsch, aber mit einem leichten ausländischen Akzent.
- B. Er spricht sehr gut Deutsch, aber hat Schwierigkeiten mit der deutschen Aussprache.
- C. Er spricht sehr gut Deutsch und seinen Akzent finde ich sehr schön.

Während die Sätze A und B auf die ähnliche Bedeutung von Akzent und Aussprachefehler hindeuten, stellt der Satz C eine andere Sichtweise des fremden Akzents dar. Hier wird der Akzent als eine interessante und neue Version der Aussprache und nicht als ein Fehler angesehen. Man könnte aber auch davon ausgehen, dass der Beurteiler im Satz C den fremden

Akzent zwar als Fehler, aber als eine interessante Fehlerform interpretiert. Aber kann ein Fehler auch schön, charmant oder interessant sein? Ein anderes Beispiel, das uns zu dieser Überlegung bringt, ist die Tatsache, dass nur eine sehr geringe Anzahl von L2-Sprechern, die Deutsch als Erwachsene gelernt haben und grammatisch fehlerlos sprechen, akzentfrei sprechen kann. Wenn wir die Termini Akzent und Aussprachefehler gleichsetzen, müssten wir davon ausgehen, dass alle L2-Sprecher phonologisch fehlerhaft sprechen.

Wie wir sehen, scheint die Abgrenzung des fremden Akzents gegenüber dem Aussprachefehler nicht eindeutig durchführbar zu sein. Die Frage, inwieweit eine Akzentscheinung als Sprachfehler angesehen wird, wirft eine Problematik der Definition der Aussprachefehler auf, auf die wir nur im Zusammenhang mit unserer Untersuchung eingehen können. Da die Bewertung der Fehlerhaftigkeit der Äußerung eines Akzentsprechers das Hauptanliegen der vorliegenden Arbeit ist, müssen wir vorher klarstellen, wie wir den Schweregrad des Akzents definieren. Hier unterscheiden wir zwischen dem leichten und dem schweren Akzent, indem wir das Kriterium Verständlichkeit bzw. Deutlichkeit als Maßstab für diese Differenzierung verwenden. Als leicht bezeichnen wir den Akzent eines L2-Sprechers, der sich möglichst an alle Ausspracheregeln der Zielsprache hält und verständlich spricht, aber dessen Aussprache nicht muttersprachlich klingt und auf seine Herkunft zurückzuführen ist. Als schwer charakterisieren wir den Akzent eines L2-Sprechers, der sich nicht an die Ausspracheregeln der Fremdsprache hält, sei es aufgrund des Transfers aus der Muttersprache oder fehlender Sprachkompetenz, und dessen Aussprache fremd und unverständlich klingt.

Vor diesem Hintergrund untersuchen wir den Schweregrad der verschiedenen phonologischen und grammatischen Fehler eines Sprechers, dessen Akzent als leicht einzustufen ist. Wir versuchen also herauszufinden, welche Aussprache- bzw. Grammatik-Fehlerarten trotz der Existenz eines kontinuierlichen leichten Akzents besonders auffallen und als schwerwiegend zu bezeichnen sind.

2.1.4. Wirkung von Sprechfehlern

„Wer sagt was mit welchen Mitteln zu wem mit welcher Wirkung?“ Die Antwort auf die von Lasswell (1952:12) aufgestellte Frage bildet die Grundformel für die Schilderung der Struktur und des Verlaufs eines Kommunikationsprozesses. Obwohl bei dieser Darstellung die

Intention des Sprechers unberücksichtigt bleibt, umfasst sie alle grundlegenden Komponenten der menschlichen Kommunikation. Ein wichtiges und für uns relevantes Element bezieht sich auf die kommunikative Wirkung, die eine sprachliche Äußerung auf den Rezipienten ausübt. Der Hörer wird in diesem Zusammenhang nicht als ein passives Wesen, das sich nur auf das Verstehen des Gesagten konzentriert, sondern als ein Akteur verstanden, der die sprachliche Äußerung inhaltlich interpretiert und in Verbindung mit dem psychologischen Eindruck, den der Sprecher auf ihn macht, auf sie reagiert.

Hirschfeld, Neuber & Stock (2010:54) unterscheiden zwischen der *direkten* und der *indirekten* Sprechwirkung. Die *direkte* Wirkung bezieht sich auf die unmittelbare Reaktion des Hörers, z. B. die Befolgung einer Aufforderung oder das Verstehen des Gesagten. Die *indirekte* Wirkung versteht sich als eine mittelbare Reaktion des Rezipienten, z. B. die Änderung der persönlichen Einstellung gegenüber dem Sprecher. Eine der wichtigsten Ursachen, die sowohl die direkte als auch die indirekte Wirkung einer sprachlichen Äußerung beeinträchtigt, ist die Abweichung von der Normsprache. Ein abweichender Sprechstil oder eine andersartige Aussprache kann sogar für L1-Sprecher Einfluss auf ihre soziale Akzeptanz in der L1-Gesellschaft ausüben.

Johansson (1973:105) unterteilt die negative Wirkung der sprachlichen Fehler auf Kommunikation in zwei Kategorien. Nach seiner Meinung können Fehler einerseits die Verständlichkeit der Äußerung beeinträchtigen, andererseits können sie den Hörer irritieren und zur Beschädigung der Sprecher-Hörer-Beziehung führen. Die Beeinträchtigung der Verständigung zeigt sich in verschiedenen Stufen und Formen, die von Kolde (1980:179-180) ausführlich dargestellt werden. Eine Kommunikation kann seiner Ansicht nach scheitern, wenn der Hörer die sprachliche Intention des Sprechers trotz seines Interesses an der Kommunikation gar nicht versteht. Manchmal kommt es aber vor, dass der Rezipient aufgrund seines Interesses am Verstehen den fehlerhaften Ausdruck entsprechend seiner eigenen Sprachnormen interpretiert und vermutet, was der Sprecher sagen will. Diese Vermutung kann insofern manchmal zur Ablenkung des Hörers vom Inhalt des Gesprächs durch die Konzentration auf seine eigene Interpretation des Gesagten und so zum Inhaltsverlust führen. Eine andere extreme Situation, die vorkommen kann, ist, dass der Rezipient die sprachliche Intention des Sprechers versteht, aber so tut, als ob er sie nicht verstanden hat. Dieser Extremfall kann nur auf fehlende soziale Akzeptanz und eine schlechte Sprecher-Hörer-Beziehung zurückgeführt werden.

Obwohl Verständigungsprobleme für den Erfolg oder das Scheitern der Kommunikation eine wesentliche Rolle spielen, darf die Wichtigkeit der psychologischen und sozialen Wirkung von Sprachfehlern nicht übersehen werden. Wir kennen alle Situationen, in denen eine fremdsprachliche Äußerung vollkommen verständlich, aber mit sprachlichen Fehlern ausgedrückt wird. Wenn wir von L1-Sprechern verlangen, diese Äußerungen hinsichtlich ihres Störungsgrads zu beurteilen, erhalten wir ganz unterschiedliche Bewertungen. Die Tatsache, dass einige Abweichungen von der Normsprache im Vergleich zu anderen gravierender auffallen, kann vor allem auf die psychologische, soziologische oder, wie oben ausgeführt, auf die *indirekte* Wirkung von Fehlern zurückgeführt werden. Die Frage, warum z. B. ein Fehler in der Positionierung des Verbs schwerwiegender zu sein scheint als ein Kasusfehler, kann nur mit rein linguistischen Mitteln schwer beantwortet werden.

Die meiste Forschung über die psychosoziale Wirkung der Fehler bezieht sich auf die abweichende Aussprache. Hirschfeld (1995:6) betont die Wichtigkeit der korrekten Aussprache bei der Bestimmung des sozialen Ansehens des Sprechers und meint: „Zu denken ist nicht nur an das Verstehen und Verstandenwerden in der Kommunikation, sondern auch an das soziale Prestige, das mit einer guten Aussprache wächst.“

Die vorliegende Untersuchung versteht sich vor allem als ein Versuch, diese *indirekte* Wirkung auf L1-Sprecher sowohl bei grammatischen als auch bei phonologischen Fehlern zu vergleichen. Der Gegenstand unserer Arbeit ist insofern die Beobachtung der psycholinguistischen Wirkung, die verschiedene Fehlerarten völlig abgesehen von Verständlichkeitsproblemen auf die L1-Rezipienten ausüben.

2.2. Forschungsstand

2.2.1. Stand der Forschung zum Thema „Integration aus sprachlicher Sicht“

Die Veröffentlichungen zum Thema „sprachliche Integration“ beschränken sich oft auf die statistischen Informationen zum Integrationserfolg der Migranten in Deutschland. Das Bundesamt für Migration und Flüchtlinge (BAMF) betreibt viel wissenschaftliche Forschung zu den Themen Migration und Integration. In einer im Rahmen des Forschungsprojekts „Integrationspanel“ durchgeführten Untersuchung, die im Jahr 2011 abgeschlossen wurde, wurde der Integrationsverlauf der Deutschkursteilnehmer mit verschiedenen Sprachniveaus beobachtet. Bei dieser Untersuchung, die durch die Befragung der Teilnehmenden während des Integrationskurses in verschiedenen Stufen sowie drei Jahre nach Beendigung des Kurses durchgeführt wurde, wurde detailliert auf den Erfolg der gesellschaftlichen Integration der Teilnehmer nach dem Absolvieren des Integrationskurses eingegangen. Die Ergebnisse dieser Studie wiesen auf einen großen Einfluss der Sprachkurse auf den Integrationsprozess der Kursteilnehmer hin. In vielen Bereichen der gesellschaftlichen Aktivitäten, z. B. beim Konsum der deutschsprachigen Medien oder beim Einstieg in den Arbeitsmarkt sowie bei der Betreuung der Kinder, zeigten die ehemaligen Integrationskursteilnehmer im Vergleich zu einer Kontrollgruppe der Migranten, die keine Sprachkurse besucht hatten, einen größeren Integrationserfolg (vgl. Lochner, Büttner & Schuller 2013).

In einer ähnlichen Studie dieses Projekts wurden die Teilnehmenden der Alphabetisierungskurse hinsichtlich ihrer Lernfortschritte sowie ihrer gesellschaftlichen Aktivitäten nach dem Besuch des Kurses befragt. Die Mehrheit der Kursteilnehmer wies bei dieser Befragung eine deutliche Verbesserung ihrer Deutschkenntnisse in allen Sprachfertigkeiten auf und bei einem großen Anteil der Teilnehmer wurde eine wesentliche Steigerung ihrer gesellschaftlichen Aktivitäten, z. B. die Zunahme ihrer Kontakte zu Deutschen oder eine Erhöhung der Anzahl ihrer Mitgliedschaften in deutschen Vereinen, festgestellt (vgl. Schuller, Lochner, Rother & Hörner 2012).

Haug (2008) berichtete in einer Studie der Reihe Integrationsreport des Bundesamtes für Migration und Flüchtlinge über die sprachliche Integration von fünf großen Migrantengruppen, nämlich den Migranten aus Italien, Griechenland, Polen, der Türkei und dem ehemaligen Jugoslawien. Laut dieser Studie fallen den Migranten die mündliche Kommunikation mit Deutschen und die aktive Beteiligung an Diskussionen schwerer als das

Hörverstehen. Die Fertigkeiten Schreiben und Lesen zeigten sich insgesamt problematischer als das Hören und Sprechen und die meisten sprachlichen Schwierigkeiten wurden bei offiziellen Angelegenheiten, z. B. bei der Kommunikation mit den Behörden, festgestellt. Nachholbedarf der Integrationskurse zeigte sich außerdem hauptsächlich bei älteren türkischen Frauen und polnischen Männern.

Wie wir sehen, richtet sich das Forschungsziel der meisten Untersuchungen zum Thema „Sprache und Integration“ auf die Bestimmung des Erfolgsgrads der Sprachkurse und ihre Rolle beim Integrationsverlauf. Eine detaillierte Überprüfung der unterschiedlichen Rolle von Grammatik, Aussprache oder anderen linguistischen Ebenen sowie verschiedenen Sprachfertigkeiten beim Integrationserfolg findet hierbei leider keine ausreichende Berücksichtigung.

2.2.2. Stand der Forschung zum Thema „Die Wirkung des ausländischen Akzents auf die Muttersprachler“

Im deutschsprachigen Raum existiert eine Reihe von Arbeiten, die sich mit der Problematik der abweichenden Aussprache der Migranten und ihre Wirkung auf deutsche Muttersprachler beschäftigen. Ursula Hirschfeld (1994) versuchte in ihrer Habilitationsschrift durch experimentelle Forschung herauszufinden, ob zwischen den phonetisch abweichenden Äußerungen der Deutschlernenden und deren Verständlichkeit für Deutsche ein Zusammenhang besteht. In ihrer Untersuchung ließ sie je eine Gruppe von deutschlernenden Studenten und nichtmuttersprachlichen Deutschlehrern Texte, Einzelsätze, Wörter und Minimalpaarwörter vorlesen. Das aufgenommene Tonband sollte sowohl von erfahrenen Phonetiklehrern als auch von naiven Hörern beurteilt werden. Dabei sollten die Probanden neben der allgemeinen Bewertung des Gehörten einige Fragen zum Textverständnis beantworten und sich bei Minimalpaarwörtern, die die wichtigsten phonetischen Merkmale der deutschen Vokale und Konsonanten beinhalteten, für die gehörte Variante entscheiden. Die Ergebnisse dieser Untersuchung deuteten darauf hin, dass die prosodisch von der Norm abweichende Aussprache nur zum Teil akzeptierbar ist und zur Beeinträchtigung des auditiven Wahrnehmungsprozesses und zu sprachlichen Missverständnissen führen kann. Außerdem stellte sich heraus, dass die Abweichungen im suprasegmentalen Bereich, insbesondere die falsche Akzentposition, die Verständlichkeit einer Aussage stärker beeinträchtigen als die segmentalen. Zudem wiesen verschiedene segmentale Abweichungen

unterschiedliche Wirkungen auf. Die vokalischen Abweichungen haben sich als Beeinträchtigungsfaktor schwerwiegender erwiesen als die konsonantischen und die Abweichungen der Vokalquantität störten die Verständlichkeit mehr als die falsche Realisierung der Vokalqualität.

Nossok (2009) orientierte sich an Hirschfelds Untersuchung und analysierte neben einem kontrastiven Vergleich der Phonologie Weißrussisch-Deutsch die Wirkung der Ausspracheabweichungen von sechs weißrussischen Sprechern auf deutsche Muttersprachler. In ihrer Untersuchung sollten 60 Probanden, bestehend aus 30 Studenten der Sprachwissenschaft und 30 der Wirtschaftswissenschaft, nach dem Hören der vorgelesenen Texte die Sprechweise der Sprecher beurteilen. Zudem sollten sie die Abweichungen, die für sie besonders kommunikationsstörend waren, durch das Ankreuzen verschiedener vorgegebener Fehlerarten benennen. Dabei wurde erwiesen, dass die Sprechweise der Sprecher mit gutem Ausspracheniveau von den Deutschen insgesamt besser beurteilt wird als die der Lernenden mit niedrigem phonetischem Niveau. Außerdem stellte sich heraus, dass die vokalischen Abweichungen den Kommunikationsprozess mehr beeinträchtigen als die konsonantischen und die falsche Aussprache der Vokalquantität störender wirkt als die fehlerhafte Realisierung der Vokalqualität. Unter den suprasegmentalen Abweichungen zeigte sich auch die falsche Position des Wortakzents als besonders störend.

Cohrs (2008) untersuchte in ihrer Dissertation den Einfluss der prosodischen Interferenzen ungarischer DaF-Lehrer auf deutsche Muttersprachler und ungarische DaF-Lernende. Die drei Sprecher dieser Studie wurden nach der phonetischen Bewertung einer Experten-Hörergruppe als Sprecher mit gutem, mittlerem und niedrigem Ausspracheniveau ausgewählt. Bewertet wurden ihre Ausspracheleistungen beim freien Sprechen, beim Vorlesen eines Textes sowie eines Kindergedichtes. Die Hörergruppe bestand aus 80 deutschen Studenten und 60 ungarischen Hörern, die in Ungarn Germanistik studierten. Die Ergebnisse der Bewertungen deutscher Hörer deuteten besonders beim gebundenen Sprechen, d. h. beim Vorlesen darauf hin, dass die abweichende Aussprache zu weniger Akzeptabilität und Verständlichkeit der Äußerung und sogar zu einer negativen Persönlichkeitsbeurteilung der Sprecher führt. Die ungarischen Hörer bewerteten dagegen die Abweichungen beim freien Sprechen negativer als beim Vorlesen des Textes und Gedichtes, bewerteten aber insgesamt die prosodischen Abweichungen positiver als die deutschen Hörer.

Eine weitere Studie, die sich mit sozialer Akzeptanz unterschiedlicher Normabweichungen der L2-Aussprache des Deutschen beschäftigt, wurde von Settinieri (2011) durchgeführt. In ihrer Untersuchung wurde die Wirkung von neun segmentalen Fehlern, die für L2-Sprecher des Französischen und des Russischen typisch waren, untersucht. Pro Fehler wurden zwei Sätze, die jeweils in zwei Varianten, nämlich einmal mit Fehlern und einmal fehlerfrei, von vier deutschen Muttersprachlern (zwei Frauen und zwei Männern) gesprochen. Die Hörergruppe bestand aus 100 deutschen Muttersprachlern (50 Frauen und 50 Männern), die die gesprochenen Sätze anhand einer siebenstufigen Skala im Hinblick auf ihre soziale Akzeptabilität bewerteten. Die Ergebnisse zeigten, dass die unterschiedlichen segmentalen Fehlertypen von L1-Hörern signifikant unterschiedlich bewertet wurden. Besonders unauffällig bzw. positiv wurden die Fehler regressive Assimilation an Stimmhaftigkeit statt progressive Assimilation an Stimmlosigkeit (*Sie findet ihn niedlich.*) und /ŋg/ statt /ŋ/ (*Er überlegt lange.*) und besonders negativ wurde der Fehler /z/ statt /ts/ (*Sie geht in den Zoo.*) bewertet. Bei dieser Untersuchung wurde außerdem der russische Akzent im Deutschen signifikant negativer bewertet als der französische.

Auch im nicht-deutschsprachigen Raum beschäftigten sich einige Linguisten mit der Problematik der Evaluation von Aussprachefähigkeiten bzw. des Sprechakzents der Migranten durch Muttersprachler. Einige Linguisten wie Anisfeld, Bogo & Lambert (1962) behandelten in ihrer Untersuchung die Wirkung der abweichenden Aussprache auf die Kontrollhörer. In ihrer Studie beobachteten sie die Reaktion der englischen Muttersprachler sowie der jüdischen Migranten auf aufgenommene Beispiele, die Englisch entweder akzentfrei oder mit jüdischem Akzent präsentierten. Bei diesem Experiment wurde die zweite Gruppe in Bezug auf bestimmte Persönlichkeitsfaktoren insgesamt schlechter beurteilt als die akzentfrei sprechende Gruppe. Die jüdische Beurteilergruppe hielt interessanterweise die akzentfreien Aufnahmen für die von Juden produzierten Beispiele, weil sie aufgrund der Vertrautheit mit diesem Akzent nicht in der Lage war, ihn herauszuhören.

Einige Linguisten wie Baird (1969, zit. nach Hellwig-Fábián 2007:33) beschäftigten sich mit der Wirkung unterschiedlicher Akzentstufen auf die Muttersprachler. Er stellte in seiner Studie heraus, dass abgesehen davon, wie stark der sogenannte *Black English* Akzent gesprochen wurde, er gleichermaßen negativ bewertet wurde. Dies wurde mit der negativen Einstellung der Hörer gegenüber den Schwarzen begründet.

Ryan, Carranza & Moffie (1977) untersuchten ebenso die Reaktion amerikanischer Muttersprachler auf Sprecher mit unterschiedlich stark ausgeprägtem spanischem Akzent. Bei dieser Untersuchung wurden sowohl die persönlichen Eigenschaften als auch die Aussprachekenntnisse der Sprecher beurteilt und im Gegensatz zu dem Untersuchungsergebnis von Baird (1969) stellte sich heraus, dass die Hörer zwischen den Akzentstufen unterscheiden und die Sprecher mit leichterem Akzent besser bewerten.

2.2.3. Stand der Forschung zum Thema „Kontrastive Beobachtung der Wirkung von Aussprache- sowie Grammatikfehlern“

Eine kontrastive Beobachtung der Wirkung von Aussprache- und Grammatikfehlern ist in einer Reihe von Studien wie in der von McLendon (1999, zit. nach Hellwig-Fábián 2007:31) durchgeführten Untersuchung zu finden. Er versucht in einer empirischen Beobachtung herauszufinden, ob russische Muttersprachler die Russischkenntnisse der Amerikaner, die Russisch mit oder ohne Aussprache- bzw. Grammatikfehlern sprechen, unterschiedlich beurteilen. In seiner Untersuchung ließ er die Sprecher jeweils einen grammatisch korrekten sowie einen fehlerhaften Text vorlesen und beobachtete dabei die Evaluation der Muttersprachler. Bei dieser Studie stellte sich heraus, dass die Aufnahmen der Sprecher mit guter Aussprache sowohl bei den grammatisch korrekten als auch bei den fehlerhaften Texten insgesamt besser beurteilt wurden als die der Sprecher mit schlechter Aussprache. Aus diesem Ergebnis schloss McLendon, dass die gute Aussprache bei der Beurteilung der Sprachkenntnisse eine wichtigere Rolle spielt als die Beherrschung der Grammatik.

Eine ähnliche Untersuchung führte Hellwig-Fábián (2007) durch. In ihrer Studie erforscht sie, wie junge Deutsche bzw. Auszubildende und Studenten Sprecher mit unterschiedlichen Aussprache- und Grammatikkenntnissen, die Deutsch mit einem ostslavischem Akzent sprechen, in Bezug auf Kriterien der Solidarität und des Status³ beurteilen. Bei dieser Untersuchung sollte ein Text, der eine Verschriftlichung der gesprochenen Antwort auf die Frage „Wohnen Sie gern in Erlangen?“ war, von demselben Sprecher einmal mit und einmal

³ Fábián lehnt sich hier an die von Brennan & Brennan (1981b) verwendeten Gegensatzpaare und übernimmt für den Faktor Status Adjektivpaare wie intelligent/dumm, gebildet/ungebildet, erfolgreich/erfolglos, wohlhabend/arm und für den Faktor Solidarität Adjektivpaare wie gut/schlecht, vertrauenswürdig/nicht vertrauenswürdig und freundlich/unfreundlich.

ohne Grammatikfehler gesprochen werden. Die Sprecher wurden nach einem Vortest in drei Akzentstufen, nämlich „sehr gut“, „gut“ und „schlecht“ verteilt. Bei dieser Untersuchung wurden unter anderem folgende Ergebnisse festgestellt: Die Einstellung junger Deutscher in Bezug auf Solidarität und Status war gegenüber Sprechern mit schlechter Aussprache nicht signifikant schlechter als gegenüber Sprechern mit guter Aussprache. Ein direkter Zusammenhang zwischen der schlechten Grammatik und einer negativen Bewertung in Bezug auf Status und Solidarität konnte auch nicht bestätigt werden. Es wurde zudem bewiesen, dass Sprecher mit Akzent selbst bei grammatisch einwandfreien Äußerungen, den Eindruck erwecken, dass sie einen Fehler machen.

Herbst (1992, zit. nach Hellwig-Fábián 2007) stellte in seiner Studie heraus, dass Sprachfehler, die von Nichtmuttersprachlern produziert werden, auffälliger und hörbarer sind als die von Muttersprachlern. In seiner Untersuchung wurden englische Texte in verschiedenen Versionen, nämlich fehlerfrei oder mit unterschiedlichen Fehlertypen, einmal von englischen Muttersprachlern und einmal von deutschen Anglistikstudenten vorgelesen. Die Muttersprachler schnitten in allen Fällen mit einer besseren Bewertung ab als die Nichtmuttersprachler. Ähnliche Ergebnisse erreichten Tucker und Sarofim (1979), die die Beurteilung der von einem englischen Muttersprachler und einem Ägypter gesprochenen Sätze, die unterschiedliche Fehlertypen enthielten, miteinander verglichen. Auch bei dieser Untersuchung wurden die vom Muttersprachler produzierten fehlerhaften Sätze als akzeptabler und weniger von der grammatischen Norm abweichend bewertet als dieselben Sätze, die vom ägyptischen Sprecher gesprochen wurden.

Raisler (1976) verwendete in ihrer Untersuchung sogar einen grammatisch korrekten Text als Untersuchungsmaterial und ließ ihn von einem englischen Muttersprachler und einem Akzentsprecher vorlesen. Interessanterweise wurden bei dem gesprochenen Beispiel vom Nichtmuttersprachler einige grammatische Fehler herausgehört, wobei sie gar nicht existierten.

2.2.4. Stand der Forschung zum Thema „Gewichtung verschiedener Sprachfehler hinsichtlich ihrer Wichtigkeit in der L1-Umgebung“

Einer der ersten Linguisten, der versucht hat, verschiedene Fehlertypen von L2-Sprechern hinsichtlich ihrer Wirkung auf deutsche Muttersprachler zu analysieren, war Politzer (1978).

In seiner Untersuchung wurden 120 deutsche Sätze, die jeweils einen Fehler enthielten, als 60 Satzpaare präsentiert. Die Sätze wurden von einer bilingualen (Deutsch-Englisch) Amerikanerin, die Deutsch mit einem leichten amerikanischen Akzent sprach, gesprochen. Insgesamt handelte es sich um sechs mögliche Fehlertypen in den Bereichen Phonologie, Kasus, Verb-Morphologie, Genus, Wortstellung und Lexik, die jeweils in vier verschiedenen Ausprägungen, die wiederum aus 5 Sätzen bestanden, präsentiert wurden. Es wurden z. B. im Bereich der Phonologie folgende Fehlertypen, die jeweils aus fünf Sätzen bestanden, berücksichtigt: falsche Realisierung des Umlauts, der Konsonanten [z] und [w], des *ich/ach*-Lautes und falsche Positionierung des Wortakzents. In dieser Untersuchung wurden alle sechs Fehlerbereiche miteinander verglichen.

Die Beurteilergruppe bestand aus 146 deutschen Schülern im Alter von 13 bis 17 Jahren, die Englisch auf einem fortgeschrittenen Niveau in der Schule lernten. Bei jedem Satzpaar sollten sich die Beurteiler für einen Satz entscheiden, in dem der begangene Fehler schwerwiegender war. Das Resultat wurde als eine Liste präsentiert, die die sechs Fehlertypen in einer Rangliste nach ihrem Wichtigkeitsgrad platzierte. Auf Platz 1 bis 6 standen: 1. Lexik 2. Verb-Morphologie 3. Wortstellung 4. Genus 5. Phonologie und 6. Kasus.

In dieser Studie wurden keine klaren Unterschiede zwischen der Wirkung von verschiedenen Fehlerarten innerhalb einer bestimmten Gruppe festgestellt. Genauer gesagt wurde z. B. nicht ausführlich analysiert, welche der eingeführten phonologischen Aspekte, nämlich die falsche Realisierung der Konsonanten, des Umlauts oder die falsche Akzentposition, am schwerwiegendsten auffiel. Es wurde allerdings auf einige interessante Ungleichheiten hingewiesen, die für unsere Studie von Bedeutung sein können, z. B. auf die unterschiedliche Wirkung von falschen Kasus-Signalen nach dem Verb und nach der Präposition, wobei die falsche Kasusendung nach dem Verb als auffälligerer Fehler markiert wurde. Außerdem wurde die falsche Position des Verbs im Hauptsatz als schwerwiegenderer Fehler bezeichnet als die fehlerhafte Verbposition im Nebensatz.

Delisle (1982) setzte die Studie von Politzer fort und führte ein ähnliches Experiment durch. In ihrer Untersuchung verwendete sie exakt dieselben Sätze wie Politzer, d. h. 120 Sätze, die jeweils einen Fehler enthielten, aber präsentierte sie als 60 geschriebene Satzpaare. Der einzige Unterschied zwischen den beiden Untersuchungsmaterialien bestand darin, dass Delisle phonologische Fehler durch Rechtschreibfehler ersetzte. Auch hier sollten sich 193

deutsche Schüler im Alter von 10 bis 17 Jahren bei jedem Paar für die Variante, die einen auffälligeren Fehler enthielt, entscheiden. In folgender Tabelle vergleicht Delisle (1982:40) ihre Ergebnisse mit der von Politzer (1978) präsentierten Rangliste.

(1) Vergleich der Ergebnisse der Studien von Politzer und Delisle

Hierarchy of Errors According to Politzer's Study (Age Ten to Seventeen)		Hierarchy of Errors According to Delisle's Study (Age Thirteen to Seventeen)	
vocabulary	77%	vocabulary	66%
verb morphology	55%	verb morphology	64%
word order	54%	gender	56%
gender	51%	word order	51%
phonology	36%	case ending	41%
case ending	28%	spelling	21%

Wie wir sehen, zeigt die oben dargestellte Tabelle eine ähnliche Ordnung bei der Klassifizierung der Fehler in beiden Untersuchungen. Die kleinen Unterschiede zwischen beiden Ergebnissen wurden auf den Unterschied zwischen der Wahrnehmung von geschriebener und gesprochener Sprache zurückgeführt.

Delisle kritisierte ihre Arbeit durch die Anmerkung, dass die Liste der Fehler, die untersucht wurden, nicht vollständig sei. Bei der Analyse der verschiedenen Ausprägungen der Fehlerkategorie *word order* wies sie genauso wie Politzer darauf hin, dass die falsche Verbposition im Hauptsatz schwerwiegender sei als im Nebensatz, wobei es zwischen der Auswertung verschiedener Konjunktionen in Nebensätzen Unterschiede zu geben scheint. Kritisch muss bei dieser Untersuchung angemerkt werden, dass die Aufgabe, sich als Muttersprachler für einen der beiden fehlerhaften Sätze entscheiden zu müssen, eine Überforderung darstellen kann, die in manchen Fällen zu ungenauen Angaben bei der Beurteilung führt. Außerdem konnte eine solche Beurteilung keine Informationen darüber geben, wie groß der Unterschied zwischen der Fehlerhaftigkeit beider gesprochenen Sätze war, denn es handelte sich nicht um eine Beurteilung durch Skalenpunkte, sondern um die Entscheidung zwischen schlecht und schlechter.

3. Vorbereitung und Durchführung der Untersuchung

3.1. Fragestellung und Hypothesen

Zur Konzeption der Ausgangshypothesen dieser Arbeit trugen verschiedene miteinander korrelierende Fragestellungen bei. Den ersten Impuls für die Formulierung der Hypothesen gab uns die Tatsache, dass die Beherrschung der Aussprache bzw. Grammatik bei der Beurteilung der Sprachkenntnisse durch Muttersprachler unterschiedliche Rollen spielen. Die Ausgangsfrage unseres Experiments lautete von daher: Abweichungen auf welcher Ebene der Sprache beeinträchtigen die sprachliche Wahrnehmung von Muttersprachlern mehr, Grammatik bzw. Morphosyntax oder Phonologie?

Eine Reihe von Linguisten wie Hirschfeld und Heike sind der Meinung, dass die Beherrschung der Ausspracheregeln einen größeren Beitrag zur erfolgreichen Kommunikation leistet als die Beherrschung der Grammatik. Heike (1972:3) verbindet die Akzeptabilität der Ausländer in einer fremden Kultur mit der Beherrschung der Aussprache und nennt sie als wichtigeren Faktor als gute Grammatikkenntnisse.

Erfahrungen mit Fremdsprachenerlernung in kommunikativen Situationen im Ausland zeigen, dass die Akzeptabilität von Ausländern, die sich die Aussprachenormen der Fremdsprache bei mangelhafter Beherrschung der Grammatik weitgehend aneignen, größer ist als im umgekehrten Fall.

Hirschfeld (1995:6) bestätigt diese Aussage und meint, dass die Beurteilung der Persönlichkeit des Sprechers von seiner phonetischen Kompetenz abhängt. Sie geht davon aus, dass das soziale Prestige durch eine gute Aussprache wächst bzw. die Persönlichkeit des Sprechers durch eine schlechte Aussprache abgewertet werden kann. Auch Adamcová (2002:79) geht bei ihrer Darstellung der Ziele von Ausspracheschulung von dem Prinzip aus, dass die falsche Aussprache zur Ablenkung des Hörers vom Inhalt der Aussage führt und so den Verständigungsprozess mehr beeinträchtigt als die grammatischen oder lexikalischen Fehler.

Wenn wir von einer unterschiedlichen Wirkung von Aussprache- und Grammatikfehlern ausgehen, stellt sich die Frage, welche Aussprache- bzw. Grammatikfehlerarten für Muttersprachler hörbarer bzw. störender und welche unauffälliger sind. Betrachten wir die Befunde von Hirschfeld (1994) und Nossok (2009) über die Wirkung der phonologischen Abweichungen auf deutsche Muttersprachler im Kapitel 2.2.2., können wir von einer

Differenzierung zwischen den segmentalen bzw. suprasegmentalen sowie den vokalischen und konsonantischen Abweichungen ausgehen. Obwohl diese zwei Studien die Wirkung von sprachlichen Fehlern aus zwei verschiedenen Perspektiven, nämlich der Beeinträchtigung des Verstehens, was im Mittelpunkt von Hirschfelds Untersuchung stand, und der Identifizierung der kommunikationsstörenden Fehlerarten, die in der Studie von Nossok angestrebt wurde, untersuchten, erzielten sie ähnliche Ergebnisse. Da die Tolerierbarkeit einer abweichenden sprachlichen Äußerung ein weiterer und wichtiger Aspekt der Wirkung von sprachlichen Fehlern ist, stellen wir uns die Frage: Kann man diese Differenzierung von Fehlertypen auf die Tolerierbarkeit der Fehler übertragen, auch wenn sie keine Verständigungs- oder Kommunikationsprobleme verursachen?

Auf der Basis dieser Beobachtungen stellten wir unsere erste Hypothese zu der unterschiedlichen Wirkung der phonologischen Abweichungen auf und gingen davon aus, dass suprasegmentale Abweichungen wichtiger als segmentale und Vokalfehler auffälliger als Konsonantenfehler sind. Wir haben außerdem angenommen, dass Vokalquantität-Fehler schwerwiegender sind als Fehler der Vokalqualität.

Auch im Bereich der Grammatik hat eine Reihe von Linguisten bewiesen, dass die Wirkung verschiedener Grammatikfehler unterschiedlich ist. Für die Aufstellung der zweiten Hypothese, die sich auf die Differenzierung zwischen den Grammatik-Fehlertypen bezieht, stützten wir uns auf die Befunde von Politzer (1978) und Delisle (1982), die die unterschiedliche Wirkung von Sprachfehlern auf eine Gruppe minderjähriger Deutscher untersucht haben. Dabei gingen wir davon aus, dass Genus-Abweichungen auffälliger sind als Kasusfehler und eine falsche Verb-Morphologie schwerwiegender als eine falsche Wortstellung. Außerdem haben wir ein weiteres Ergebnis ihrer Studie, dass eine fehlerhafte Verbposition im Hauptsatz hörbarer ist als eine falsche Verbposition im Nebensatz, in die Hypothese 2 aufgenommen.

Für die Formulierung der Hypothese 3 stützen wir uns auf die Ausgangsfrage unserer Untersuchung und setzen einen kontrastiven Vergleich der Wirkung von Aussprache- und Grammatikfehlern in den Mittelpunkt dieser Hypothese. Hier wurde angenommen, dass die Aussprachefehler der DaF-Sprecher von deutschen Muttersprachlern weniger toleriert werden als ihre grammatischen Fehler.

Wenn wir von einer unterschiedlichen Wirkung verschiedener Sprachfehler, die in den oben angeführten Hypothesen vorausgesetzt sind, ausgehen, stellt sich die Frage, welche Fehlerarten bewusster und sicherer erkannt werden. Angesichts der Tatsache, dass für die phonetisch ungeschulten Hörer wie die Mehrzahl unserer Probanden, die wir später im Kapitel 3.3.3. ausführlich vorstellen werden, die Erkennung der grammatischen Fehler aufgrund ihrer erworbenen Vorkenntnisse in der Schule einfacher ist als die Feststellung der phonologischen Fehler, gingen wir in der Hypothese 4 davon aus, dass Grammatikfehler einfacher bzw. sicherer erkannt werden als Aussprachefehler.

Vergleichen wir die Hypothesen 3 und 4 miteinander, können wir eine Unstimmigkeit zwischen beiden Annahmen feststellen. Während wir in der Hypothese 3 vermuten, dass die Grammatikfehler von deutschen Hörern als weniger störend bewertet werden, gehen wir in der Hypothese 4 davon aus, dass die fehlerhafte Grammatikform leichter erkannt wird als die Aussprachefehler. Um diese Diskrepanz zu durchleuchten, stellten wir eine weitere Hypothese zur allgemeinen Wahrnehmung der Sprechfehler auf und gingen davon aus, dass zwischen der richtigen Fehlererkennung und dem Störungsgrad der Fehler ein Zusammenhang besteht.

In Bezug auf die allgemeine Fehlerwahrnehmung stellten wir außerdem unsere letzte Hypothese auf, in der vermutet wird, dass ein fremder Akzent zu einer falschen Bewertung der Sprachkenntnisse führen kann. Solche falschen Beurteilungen, dass ein fehlerfreier Text als fehlerhaft beurteilt wurde, weil er von einem Nichtmuttersprachler gesprochen wurde, konnten, wie im Kapitel 2.2.3. erwähnt, in einigen linguistischen Studien wie in der Untersuchung von Raisler (1976) bewiesen werden.

An dieser Stelle fassen wir die aufgestellten Hypothesen der folgenden Studie wie folgt zusammen:

Hypothesen zu den phonologischen Fehlerarten

Hypothese 1.1: Suprasegmentale Sprechfehler sind auffälliger als segmentale Fehler.

Hypothese 1.2: Sprechfehler in der Realisierung der Vokale wirken störender als die der Konsonanten.

Hypothese 1.3: Eine fehlerhafte Aussprache der Vokallänge fällt schwerwiegender aus als eine falsche Realisierung der Vokalqualität.

Hypothesen zu den syntaktischen und morphosyntaktischen Fehlerarten

Hypothese 2.1: Genusfehler sind schwerwiegender als Kasusfehler.

Hypothese 2.2: Verbmorphologie-Fehler werden weniger toleriert als Verbposition-Fehler.

Hypothese 2.3: Eine fehlerhafte Verbposition im Hauptsatz ist auffälliger als eine falsche Verbposition im Nebensatz.

Hypothesen zum kontrastiven Vergleich der Grammatik- und Aussprachefehler

Hypothese 3: Insgesamt werden Sätze mit manipuliert eingeplanten Aussprachefehlern mehr als Sprachfehler wahrgenommen als diejenigen, die einen Grammatikfehler enthalten.

Hypothese 4: Grammatische Fehler sind besser bzw. einfacher (Hypothese 4.1) und sicherer (Hypothese 4.2) zu erkennen als Aussprachefehler, d. h. die abweichende Aussprache führt mehr noch als die fehlerhafte Grammatik zu einer falschen Beurteilung der Sätze bzw. zu einer falschen Erkennung des Fehlers.

Hypothesen zur allgemeinen Wahrnehmung der Sprechfehler

Hypothese 5: Es gibt zwischen der richtigen Erkennung des Fehlers im Satz und dessen negativer Bewertung eine positive Korrelation.

Hypothese 6: Sprecher mit Akzent erwecken selbst bei korrekten Sätzen den Eindruck, sie machen einen Fehler.

3.2. Auswahl der Fehlertypen

Da wir die zwei gegenüberliegenden Teilgebiete dieser Untersuchung, nämlich die Grammatikfehler und die Ausspracheabweichungen, nicht als ein abstraktes Konzept, sondern als Realisierung in verschiedenen Fehlerarten betrachten wollen, fokussieren wir uns auf konkrete grammatische sowie phonologische Merkmale. Für die Auswahl der zu untersuchenden Grammatikkriterien bezogen wir uns auf die folgenden sechs Kriterien, die nach meiner eigenen Erfahrung als DaF-Lehrende am häufigsten vorkommen und für die Lerner am wichtigsten zu sein scheinen.

(2) Die zu untersuchenden grammatischen Abweichungen

Gruppe	Fehlertyp	Beispiel
G	Genus (Artikel)	<i>ein zweites</i> statt <i>eine zweite Sprache</i>
H	Kasus	<i>Wir sind auf die</i> statt <i>auf der Straße</i>
I	Präposition	<i>Angst gegen</i> statt <i>vor</i>
J	Partizip II	<i>bereiten</i> statt <i>bereitet</i>
K	Verbposition im Hauptsatz (Wortstellung)	<i>letztes Jahr ich habe</i> statt <i>habe ich</i>
L	Verbposition im Nebensatz (Wortstellung)	<i>Obwohl er ist nicht groß</i> statt <i>nicht groß ist</i>

Bei der Auswahl der in dieser Arbeit zu untersuchenden phonologischen Abweichungen beziehen wir uns auf die Ergebnisse der im Rahmen meiner Masterarbeit (Pazan 2011) durchgeführten empirischen Untersuchung zu prosodischen Fehlern der persischsprachigen Deutschlernenden. Dabei wurden sechs Ausspracheprobleme der persischen Deutschlerner einbezogen:

(3) Die zu untersuchenden phonologischen Abweichungen

Gruppe	Fehlertyp	Beispiel
A	Realisierung der Affrikate [ts]	[sait] statt [tsait]
B	Auslautverhärtung	[na i:v] statt [na i:f]
C	Vokalquantität [a] vs. [a:] und [e:] vs. [ɛ]	[a tɐm] statt [a: tɐm]
D	Realisierung des Umlauts [y:] vs. [u:] und [ɥ] vs. [ʊ]	[u:bl] statt [y:bl]
E	Konsonantenhäufung in der Wortanlautsilbe	[ˈɛjpra:xə] statt [ˈjpra:xə]
F	Interferenzfehler bei der Positionierung des Wortakzents	Spiel platz statt Spiel platz

Obwohl in der im Rahmen meiner Masterarbeit durchgeführten Untersuchung die Realisierung der Vokale und Konsonanten nicht direkt als Hypothese für eine transferbedingte Sprachabweichung behandelt wurde, haben die deutschen Hörer bei der

Beurteilung des Zusammenhangs zwischen dem Wortakzent und der Vokalquantität die Anwendung falscher Vokale bzw. Konsonanten als besonders störend bezeichnet. Bei vielen Sprechern wurde z. B. die mangelhafte Lippenbewegung als auslösender Faktor für die abweichende Artikulation angemerkt.

Da in dieser Arbeit keine konkreten Angaben über die als fehlerhaft ausgesprochenen Vokale und Konsonanten gemacht wurden, bezogen wir uns für die Auswahl der Stimuli für diese zwei Gruppen auf die kontrastive Beobachtung der Konsonanten- sowie Vokalsysteme beider Sprachen, die wir in den nächsten zwei Kapiteln ausführlich präsentieren.

Die zwei anderen Kriterien, nämlich das Problem der persischen Sprecher beim Aussprechen der Konsonantenhäufung in der Wortanlautsilbe und die falsche Positionierung des Wortakzents, gehörten zu den Haupthypothesen der Untersuchung, die falsifiziert wurden. Wir haben diese Kriterien in unserer Arbeit aus dem Grund mitberücksichtigt, da die Bestätigung oder Ablehnung einer Hypothese in dieser Arbeit durch das Erreichen oder Nichterreichen einer subjektiven Mittelwertanzahl von 50% bestimmt wurde. Die Hypothese „Konsonantenhäufung“ wurde z. B. nicht bestätigt, weil die Durchschnittsprozentszahl der als abweichend gehörten Aussprachevarianten 28% betrug. Die Abweichungsquote bei der Aussprache einiger Wörter wie *Straße* oder *Schlange* wurde allerdings mit einer Anzahl von über 40% registriert. Der gleiche Prozess war bei der Hypothese zum Wortakzent zu beobachten. Bei einigen Wörtern wie bei *Spielplatz* zeigte sich sogar eine Abweichungsquote von 42%, was im Rahmen dieser Arbeit als falsifiziert bewertet wurde.

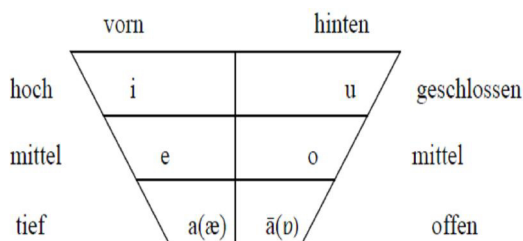
Ein weiteres konsonantisches Phänomen, das wir in dieser Studie mitberücksichtigt haben, war die Auslautverhärtung. Da es im Persischen keine Regel für den Wechsel zwischen stimmhaften und stimmlosen Versionen der Plosive und Frikative am Wortende gibt, wird erwartet, dass die persischen DaF-Sprecher die Regel der Auslautverhärtung im Deutschen nicht einhalten und die stimmhaften Plosive und Frikative auch am Wortende stimmhaft aussprechen.

3.2.1. Auswahlkriterien der Beispiele für Vokalfehler

Bei der Auswahl der Beispielwörter, die eine falsche Realisierung der Vokale enthielten, bezogen wir uns auf die kontrastive Gegenüberstellung der Vokalsysteme beider Sprachen,

die ich im Rahmen meiner Bachelorarbeit (Pazan 2008) präsentiert habe. Laut dieser Untersuchung weist im Persischen die Vokallänge kein distinktives Merkmal auf und besteht aus den Kurzvokalen [e], [æ], [o], den Langvokalen [u], [i], [ā] sowie den Diphthongen [ei] und [ou].

(4) Die Vokalphoneme des Persischen (Hamidi 2006:10)



Das deutsche Vokalsystem verfügt im Gegensatz zum Persischen über eine große Anzahl von Vokalen. Kürschners (2008:40) Darstellung nach besitzt die deutsche Sprache 17 Vokale.

(5) Die Vokalphoneme des Deutschen (Kürschners 2008:40)

		Zungenstellung			
		vorn		neutral	hinten
Zungenhöhe	hoch	i: ɪ	y: ʏ		u: ʊ
	mittel	e: ɛ	ø: œ	ə	o: ɔ
	gehoben	ɛ: ɛ		e	
	tief			a: a	
		ungerundet	gerundet	ungerundet	gerundet
		Lippenform			

Die Gegenüberstellung der Vokalsysteme beider Sprachen weist auf einige wichtige vokalische Unterschiede hin. Da die Vokallänge im Persischen kein distinktives Merkmal aufweist, kann man davon ausgehen, dass die persischen Deutschlernenden beim Sprechen kaum auf die Unterschiede zwischen langen und kurzen Vokalen achten und so fehlerhafte Vokale produzieren. Für das Kriterium Vokalquantität beschränken wir uns bei folgender Untersuchung auf die Vokale [a] und [e]. Der Grund für diese Auswahl war, dass sich bei

dem Vokal [a] die lange und die kurze Variante hinsichtlich ihrer Qualität wenig voneinander unterscheiden.⁴

Obwohl der Buchstabe „e“ durch vier Phoneme, nämlich /ɛ/, /ɛ:/, /e:/ und /ə/, realisiert wird (vgl. Wiese 2011:48), haben wir bei der Aussprache dieser Phoneme streng darauf geachtet, dass die kurze und die lange Variante in Hinblick auf ihre Qualität wenig voneinander abweichen. Die Vokale [u] und [o] wurden für dieses Kriterium nicht mitberücksichtigt, weil die beiden auch eine umgelautete Variante besitzen, die es im Persischen nicht gibt, und so von persischen Sprechern miteinander verwechselt werden könnten.

Bei der Auswahl der Beispiele für die falsche Produktion der Vokale hinsichtlich ihrer Qualität bezogen wir uns vor allem auf deutsche Vokale, die im Persischen nicht existieren. Wie in den oben aufgeführten Darstellungen zu sehen ist, verfügt das Persische über eine deutlich geringere Anzahl an Vokalen. Die Zentralvokale des Deutschen bzw. der Vokal Schwa [ə] und das vokalische [ɤ] bzw. [ɐ] finden z. B. im Persischen keine Entsprechung. Den Laut Schwa [ə] haben wir nicht als Beispiel genommen, weil die Aussprache dieses Lautes für persische Sprecher kein Problem zu sein scheint. So können sie sich durch das Ersetzen des Schwas durch das persische [e] deutlich und problemlos ausdrücken.

Der Konsonant [r] am Silbenende nach einem Vokal wird von persischen Sprechern in den meisten Fällen nicht wie ein schwaches vokalisiertes [ɐ], sondern genauso wie in der Schrift bzw. wie ein [ər] ausgesprochen. Das Wort *weiter* wird z. B. von persischen Muttersprachlern oft als [vāytər] gesprochen. In unsere Untersuchung haben wir das vokalische [ɐ] trotzdem nicht aufgenommen, weil selbst im Deutschen das konsonantische [r] in Gebieten wie in Österreich oder Bayern häufig auch in der finalen Wortposition erhalten bleibt (vgl. Kohler 1995:165).

Als wichtigste fehlende deutsche Vokalalternation im persischen Vokalsystem kann auf die Frontierung des Vokals (Umlaut) hingewiesen werden. Man kann also davon ausgehen, dass die persischen Deutschlerner bei der Aussprache dieser Vokale meistens einen Transferfehler begehen und die Umlaute [ö], [ü] und [ä] durch die persischen Vokale [o], [u] und [e]

⁴ Über die Qualitätsunterschiede zwischen /a/ und /a:/ gibt es unterschiedliche Meinungen. Einige Autoren sind der Ansicht, dass sich die beiden Varianten in der Qualität voneinander unterscheiden und bezeichnen /a/ als vorderen kurzen und /a:/ als hinteren langen Vokal. Manche andere Linguisten wie Wiese (2011:37) halten diese Qualitätsunterscheidung für unbegründet und bezeichnen beide als zentrale und tiefe Vokale.

ersetzen. Für die folgende Untersuchung haben wir uns auf die kurze und die lange Realisierung des Umlauts [ü] beschränkt, weil die Produktion der umgelauteten Form des geschlossenen [u] für die persischen DaF-Sprecher angesichts der Existenz des langen Vokals [u] im Persischen einfacher zu sprechen sein scheint. So konnte unser Sprecher zwei Versionen dieses Umlauts, nämlich einmal in korrekter und einmal in einer abweichenden Form, problemlos aussprechen. Die Behauptung, dass die Produktion des Umlauts [ü] für persische Muttersprachler einfacher wäre, bedarf allerdings einer weiteren Studie, die wir im Rahmen dieser Arbeit aus Umfangsgründen nicht berücksichtigen konnten.

3.2.2. Auswahlkriterien der Beispiele für Konsonantenfehler

Für die Zusammenstellung der Beispielsätze, die eine falsche Realisierung der Konsonanten reflektieren sollen, stützten wir uns auf die konsonantischen Unterschiede zwischen beiden Sprachen. Um beide Konsonantensysteme parallel miteinander vergleichen zu können, präsentierten wir die Konsonanten beider Sprachen in der Tabelle (6). Bei dieser kontrastiven Abbildung bezogen wir uns für das Persische auf die Darstellung von Hamidi (2006) und für das Deutsche auf die Versionen von Bußmann (2008) und Ramers (1998).

Durch die Betrachtung dieser Tabelle können wir feststellen, dass einige Konsonanten des Deutschen sich in Bezug auf Artikulationsart und -ort vom Persischen unterscheiden. Die Aussprachefehler, die sich auf die kleinen Unterschiede zwischen dem Artikulationsort bestimmter Konsonanten wie [t], [d], [k], [g], [x] oder [r] beziehen, können zwar die richtige Aussprache beeinträchtigen, der Unterschied ist aber so minimal, dass er in keiner Weise zur Sanktionierung der Sprachproduktion als fehlerhaft führen kann.

Im Deutschen existieren aber einige Konsonanten, die im Persischen keine Entsprechung finden. Das Aussprechen solcher Konsonanten, die wir in der Tabelle mit grauer Farbe markiert haben, bereitet den persischen Deutschlernern oft Schwierigkeiten. Besonders schwierig auszusprechen sind die Konsonanten [ç], uvulare [R] und Affrikaten [pf] und [ts]. Obwohl es im Persischen keine Entsprechung für die Konsonantenverbindungen /ks/ und /ŋ/ gibt, können Perser diese Phonemkombinationen in den meisten Fällen problemlos aussprechen. Die Konsonanten [ç] und [r] haben wir aus dem Grund nicht in unsere Studie aufgenommen, weil sie auch im deutschsprachigen Raum regionale Varianten aufweisen. Da die Untersuchung der Wirkung der abweichenden Aussprache all dieser Konsonanten aus

Umfangsgründen nicht möglich war, beschränkten wir uns hier auf die abweichende Aussprache der Affrikate [ts].

(6) Das Konsonantensystem des Deutschen und des Persischen⁵:

Artikulationsart	Artikulationsort								
	bilabial	labiodental	dental	dental- alveolar	palato- alveolar	palatal	velar	uvular	glottal
Stl. Plosive	p p		t	t		K	k		ʔ ʔ
Sth. Plosive	b b		d	d		g	g	q	
Stl. Frikative		f f		s s	ʃ ʃ	ç	x	x	h h
Sth. Frikative		v v		z z	ʒ ʒ	j		ʁ	
Stl. Affrikaten	pf			ts	tʃ		ks		
Sth. Affrikaten					dʒ				
Nasale	m m			n n			ŋ		
Laterale				l l					
Vibranten		r r						R	

⁵ Die fett markierten Konsonanten präsentieren die deutschen Konsonanten.

3.3. Design der Untersuchung

3.3.1. Untersuchungsmaterialien

Für die Auswahl der Testmaterialien haben wir auf ein freies Gespräch verzichtet, weil die Beobachtung der Wirkung von bestimmten bzw. manipulierten Sprechfehlern auf deutsche Muttersprachler im Mittelpunkt unserer Untersuchung stand. Die Testmaterialien in unserer Arbeit bestanden aus 36 Sätzen, die jeweils in drei Varianten dargeboten wurden. Die erste Variante präsentierte eine sowohl phonologisch als auch morphosyntaktisch korrekte Form des Satzes. Die zweite Variante enthielt einen manipulierten Aussprachefehler, wobei die dritte Version einen manipulierten Grammatikfehler enthalten sollte (s. Anhang 1). Obwohl bei der Konzipierung der Sätze darauf geachtet wurde, dass die Sätze ungefähr die gleiche Länge haben, haben wir bewusst auf gleich strukturierte Sätze verzichtet, denn diese Strukturähnlichkeit hätte zur Monotonie der gesprochenen Sätze führen und so die Bewertung indirekt beeinflussen können.

Wir haben uns hier für Sätze und nicht für einen Text entschieden, weil die Bewertung der einzelnen manipulierten Fehler in einem zusammenhängenden Text für die Hörer eine fast unmögliche Aufgabe gewesen wäre. Außerdem war die Produktion der bestimmten Aussprachefehler bei isolierten Sätzen besser zu beobachten als bei einem Text, der an bestimmten Stellen manipulierte Fehler enthalten sollte. Auch einzelne Wörter als Stimuli wurden ausgeschlossen, weil die Beurteilung der Sprachfähigkeit eines Sprechers nur durch die Bewertung des gesprochenen Wortes kaum realisierbar wäre und so hätten außerdem keine Grammatikfehler mitberücksichtigt werden können.

Bei den Sprachfehlern im Bereich der Phonologie bezogen wir uns auf diejenigen Fehler, die einerseits für persische Sprecher kontrastiv bedingt vorkommen, andererseits im deutschsprachigen Raum keine unterschiedlichen regionalen Varianten vorweisen. Auf der phonemischen Ebene lag der Schwerpunkt auf der fehlerhaften Anwendung von Allophonen. Genauer formuliert wurde hier streng darauf geachtet, dass die fehlerhafte sowie korrekte Version keine Minimalpaare bildeten, denn in so einem Fall hätten die Hörer den gesprochenen Fehler als semantischen und nicht als phonologischen Fehler verstehen können. So konnte also verhindert werden, dass die manipulierten Fehler zur Verletzung der Verständlichkeit der sprachlichen Äußerung führen.

3.3.2. Sprecher

Um Störfaktoren wie individuelle Sprechweise bzw. Stimmqualität, Sprechgeschwindigkeit, Intonation sowie geschlechtsspezifische Stimmwirkung auszublenden, bezogen wir uns auf die von Lambert, Hodgson, Gardner & Fillenbaum (1960) entwickelte Matched-Guise-Technik. Das wichtigste Merkmal dieser Methode ist, dass derselbe Sprecher einen Text oder Satz oder ein Wort in zwei Varianten, einmal in der normalen und einmal in der manipulierten Form vorträgt, d. h. ein Sprecher produziert zwei Stimuli, wobei die Beurteiler nicht darüber informiert werden, dass es sich bei den beiden um dieselbe Stimme handelt. Eine wichtige Voraussetzung für das Verwenden dieser Technik ist, dass die Sprecher die Fähigkeit besitzen, die manipulierte Version so gut vorzutäuschen, dass die Hörer diese als eine natürliche Abweichung annehmen.

Um dies zu erreichen, wurde ein bilingualer (Deutsch-Persisch) Sprecher ausgewählt, dessen Muttersprache Persisch war und der Deutsch mit einem kontinuierlich leichten Akzent sprach. Zudem ging sein Sprachniveau über das Mittelstufenniveau hinaus und lag auf dem Niveau C1-C2. Wären die Sätze stattdessen von einem Muttersprachler gesprochen worden, wäre zwar die interne Validität der Untersuchung gestiegen (vgl. Albert & Marx 2014:31), weil so die Gefahr, dass die Hörer einen anderen Aussprachefehler anstelle des gemeinten Fehlers beurteilen, hätte verringert werden können. Die Nachahmung der L2-Sprachfehler durch einen Muttersprachler hätte aber andererseits zur Reduzierung der Natürlichkeit der gesprochenen Sätze führen können, denn so hätten die eingebauten Fehler künstlich wirken können. Dies kann man auch bei den Anmerkungen von Hörern in der Untersuchung von Settinieri (2011:76), in der die Sprecher deutsche Muttersprachler waren, beobachten.

„da es sich um gestellte aufnahmen handelt war es schwieriger sie zu bewerten als wenn akzente einem in der realität begegnen [sic!]"

„die akzente klangen nicht echt, eher wie von deutschen extra nachgesprochen. das hat bei mir den effekt abgeschwächt, glaube ich [sic!]"

Entsprechend der Matched-Guise-Technik wurde in dieser Untersuchung jeder Satz einmal in der korrekten Form, die der Standardsprache entsprach, und zweimal jeweils mit einer manipulierten Abweichung, einmal in der Grammatik und einmal in der Aussprache von demselben Sprecher gesprochen. Da in jedem Satz nur ein einziger manipulierter Fehler produziert werden sollte, wurde darauf geachtet, dass der Sprecher möglichst keinen anderen

Aussprachefehler beging, was allerdings in Hinsicht auf den nichtmuttersprachlichen Akzent des Sprechers nicht hundertprozentig zu kontrollieren war.

3.3.3. Die Wahl der Stichprobe

Die Hörergruppe bestand aus 32 Probanden, bestehend aus 25 Studenten der Philipps-Universität Marburg sowie sieben berufstätigen Akademikern verschiedener Fachbereiche. Bei der Suche nach Probanden wurde streng darauf geachtet, dass die Hörer monolinguale Muttersprachler sind bzw. nicht mehrsprachig aufgewachsen. An dieser Untersuchung haben sich mehr männliche als weibliche Probanden beteiligt und die überwiegende Anzahl befand sich im Alter von 18 bis 30 Jahren. Die folgende Tabelle präsentiert die Häufigkeitsinformationen hinsichtlich des Geschlechts und Alters sowie der Fachbereichszugehörigkeit der Versuchspersonen. Detaillierte Informationen zu den einzelnen Personen finden Sie im Anhang 2.

(7) Angaben zu der Hörergruppe

Geschlecht		
	Häufigkeit	Prozent
Männlich	20	62,5
Weiblich	12	37,5
Alter		
18-30 Jahre	26	81,3
über 30 Jahre	6	18,8
Fachbereich		
FB 2: Wirtschaftswissenschaften	1	3,1
FB 3: Gesellschaftswissenschaften und Philosophie	12	37,5
FB 6: Geschichte und Kulturwissenschaften	6	18,8
FB 9: Germanistik und Kunstwissenschaften	5	15,6
FB 10: Fremdsprachliche Philologien	8	25,0
FB 12: Mathematik und Informatik	1	3,1
FB 16: Pharmazie	1	3,1
FB 20: Medizin	2	6,3
FB 21: Erziehungswissenschaften	4	12,5

3.3.4. Untersuchungsbeschreibung

Nach einer kleinen Vorbereitung wurde von dem Sprecher verlangt, die Satzpaare auf Tonband zu sprechen. Um Hintergrundgeräusche zu vermeiden und eine bessere Aufnahmequalität zu erreichen, wurde die Tonaufnahme in der Aufnahmekabine des Phonetik-Labors des Fachbereichs „Germanistische Sprachwissenschaft“ der Universität Marburg durchgeführt.

Das Beurteilungsformular bestand aus zwei Teilen. Im ersten Teil wurden die biographischen Daten der Teilnehmer, z. B. Geschlecht, Altersgruppe, Studium und Beruf, abgefragt. Im zweiten Teil wurde für die Beurteilung der gesprochenen Sätze eine Sechs-Punkte-Skala vorgelegt, deren Notenstufen zwischen sehr gut (1) und sehr schlecht (6) variierten. Hier sollten die Hörer für jeden Satz die Frage „Wie gut wurde der Satz auf Standarddeutsch gesprochen?“ durch das Ankreuzen einer Note zwischen 1 und 6 beantworten. Für jeden Satz sollten sie zudem, falls sie den Satz als fehlerhaft bewertet hatten, anmerken, welchen Fehler sie gehört haben und in einer Vier-Punkte-Skala ankreuzen, wie sicher sie sich bei dieser Erkennung waren (s. Anhang 3).

Die Verwendung der endpunktbenannten Skalen sollte uns ermöglichen, die Auswertungsmöglichkeiten der intervallskalierten Daten anzuwenden. Diese Skalen werden im Gegensatz zu verbalisierten Skalen als intervallskaliert betrachtet, denn der Abstand zwischen den Skalenpunkten wird gerne als gleich angesehen, auch wenn sie nicht hundertprozentig gleichabständig sind (vgl. Porst 2014:75). Durch die gerade Anzahl der Skalenpunkte haben wir außerdem versucht, die Bewertenden dazu zu führen, eine Entscheidung zwischen gut und schlecht zu treffen bzw. keinen Mittelpunktswert, der meist als Meinungslosigkeit interpretiert wird, anzukreuzen.

Die 36 Sätze wurden jeweils in drei Varianten, einmal in fehlerfreier Form, einmal mit einem manipulierten grammatischen Fehler und einmal mit einem manipulierten phonologischen Fehler in nach dem Zufallsprinzip ausgewählter Reihenfolge präsentiert. Um Reihenfolgeeffekt bei der Bewertung völlig zu umgehen, wurden zudem die gemischten Sätze in vier Gruppen geteilt, die für vier Gruppen von Probanden in vier verschiedenen Reihenfolgen präsentiert wurden.

Um die Probanden auf die Befragung vorzubereiten, wurde ihnen vorher erklärt, dass sie insgesamt 108 Sätze hinsichtlich der Akzeptabilität nach Kriterien der deutschen Standardsprache beurteilen sollen. Hier wurde darauf eingegangen, dass manche Sätze sich in einer zufälligen Reihenfolge wiederholen und in manchen Grammatik- sowie Aussprachefehler enthalten sind. Vor der Beurteilung wurde auf diese Weise den Probanden bewusst, dass es sich darum handelt, ob sie die gesprochenen Fehler erkennen, wie störend diese Fehler für sie klingen und wie sicher sie sich bei der Erkennung dieser Fehler sind.

Die meisten Versuche wurden im Foyer der Universitätsbibliothek in Marburg durchgeführt. Die Befragung erfolgte schriftlich und die Audiodateien bzw. Sätze wurden anhand eines Laptops und über einen Kopfhörer eingespielt. So konnten die Probanden die Lautstärke und die Pausenlänge zwischen den Sätzen selbst bestimmen. Es wurde aber darauf hingewiesen, dass sie die Sätze auf keinen Fall wiederholen dürfen. Nach der Befragung erhielten die Teilnehmer eine kleine Aufwandsentschädigung.

3.3.5. Vorbereitung auf die statistische Auswertung

Umschreibung der Hypothesen in eine statistische Notation

Um unsere Hypothesen anhand statistischer Mitteln besser prüfen zu können, müssen wir sie als präzise Aussagen formulieren, in denen Vergleiche bzw. Zusammenhänge zwischen mindestens zwei Variablen in eine statistische Notation umgeschrieben sind. Für eine statistische Hypothesenprüfung⁶ hätten wir, wie es in anderen wissenschaftlichen Bereichen üblich ist, die sogenannten Null- und Alternativhypothesen aufstellen können⁷ (vgl. Bortz & Schuster 2010:98). Auf die Formulierung der Nullhypothesen in dieser Arbeit haben wir jedoch verzichtet, weil die Aufstellung solcher Hypothesen in der Sprachlehrforschung eher unüblich ist (vgl. Albert & Marx 2014:25). Im Folgenden präsentieren wir die statistische Formulierung unserer Hypothesen.

⁶ „Ein statistischer Test ist eine Regel, die es erlaubt, für jedes Stichprobenergebnis eine Entscheidung zwischen der Null- und der Alternativhypothese zu treffen.“ (Bortz & Schuster 2010:99)

⁷ Alternativhypothesen (H_1) sind die von uns innovativ aufgestellten Hypothesen, die anhand eines Experiments oder einer Befragung bestätigt bzw. verifiziert werden. Nullhypothesen (H_0) stehen im Gegensatz zu den Alternativhypothesen und behaupten, dass der Vergleich oder Zusammenhang, deren Vorhandensein die Alternativhypothese zu bestätigen versucht, nicht existiert (vgl. Bortz & Schuster 2010:98).

Hypothese 1.1

Die suprasegmentalen Sprechfehler erzielen bei der Bewertung höhere⁸ bzw. schlechtere Mittelwerte als die segmentalen:⁹ $M(C+E+F) > M(A+B+D)$

Hypothese 1.2

Die durchschnittliche Bewertung der Gruppen mit Fehlern in der Realisierung der Vokale (Gruppen C+D) ist schlechter als die der Gruppen mit Konsonantenfehlern (Gruppen A+B): $M(C+D) > M(A+B)$

Hypothese 1.3

Der Bewertungsmittelwert der Gruppe mit Fehlern in der Vokallänge (Gruppe C) ist höher als die durchschnittliche Bewertung der Gruppe mit falscher Realisierung des Umlauts (Gruppe D): $M(C) > M(D)$

Hypothese 2.1

Die Gruppe mit Genusfehlern (Gruppe G) erzielt bei der Bewertung einen schlechteren Mittelwert als die Gruppe mit Kasusfehlern (Gruppe H): $M(G) > M(H)$

Hypothese 2.2

Verbmorphologie-Fehler (Gruppe J) werden im Durchschnitt schlechter bewertet als falsche Verbpositionen im Hauptsatz (Gruppe K) und im Nebensatz (Gruppe L): $M(J) > M(K) \ \& \ M(J) > M(L)$

Hypothese 2.3

Eine falsche Verbposition im Hauptsatz erzielt bei der Bewertung einen höheren Mittelwert als eine fehlerhafte Verbposition im Nebensatz: $M(K) > M(L)$

⁸ Höhere Werte entsprechen hier einer schlechteren Evaluation, da in unserer Rating-Skala der Wert 1 als sehr gut und der Wert 6 als sehr schlecht definiert wurden.

⁹ Als suprasegmentale Abweichung wurden drei Gruppen mit Fehlern in der Vokalquantität (Gruppe C), Konsonantenhäufung in der Wortanlautsilbe (Gruppe E) und Wortakzent (Gruppe F) ausgewählt, bei denen es Merkmale gab, die über ein Lautsegment hinausgingen (Vokalquantität), die die interne Silbenstruktur änderten (Konsonantenhäufung) und die eine abweichende Akzentposition aufwiesen (Wortakzent).

Die Gruppe der segmentalen Abweichung bestand aus drei Gruppen mit falscher Realisierung der Affrikate [ts] (Gruppe A), der Verletzung der Regel Auslautverhärtung (Gruppe B) und fehlerhafter Aussprache des Umlauts (Gruppe D).

Hypothese 3

Phonologische Fehler werden durchschnittlich schlechter beurteilt, d. h. erzielen einen höheren Mittelwert als Grammatikfehler: $M(A+B+C+D+E+F) > M(G+H+I+J+K+L)$

Hypothese 4

4.1. Richtige Erkennung des Fehlers

Grammatikfehler werden im Durchschnitt besser erkannt, d. h. erzielen eine höhere Anzahl der richtigen Erkennung als Aussprachefehler: $\sum(G+H+I+J+K+L) > \sum(A+B+C+D+E+F)$

4.2. Sicherheitsgrad bei der Erkennung

Der Mittelwert der Sicherheitswerte von Gruppen mit Grammatikfehlern ist niedriger als der von Gruppen mit Aussprachefehlern, wobei die niedrigere Zahl hier einen höheren Sicherheitsgrad aufweist¹⁰: $M(G+H+I+J+K+L) < M(A+B+C+D+E+F)$

Hypothese 5

Je höher die durchschnittliche Anzahl der richtigen Fehlererkennung, desto höher bzw. schlechter ist die durchschnittliche Bewertung. Der Korrelationskoeffizient r sollte hier im besten Fall +1 betragen, wobei unsere Hypothese so definiert wurde: $r > 0$

Hypothese 6

Die gesprochenen korrekten Sätze erzielen bei der Bewertung Werte, deren Durchschnitt größer ist als 1 (sehr gut): $M > 1$

Vorbereitung der Datensätze auf die Auswertung

Für die Analyse der Ergebnisse dieser Studie haben wir zuerst die erhobenen Roh-Daten, die in 32 einzelnen Fragebögen schriftlich notiert worden waren, in eine Excel-Tabelle übertragen und dann in das Statistik-Programm SPSS importiert. Bei dieser Eingabe wurden die Bewertungen und Äußerungen der Probanden zu den erkannten Fehlern buchstabengetreu wiedergegeben.

Einige Beurteilungen, bei denen der korrekte Satz schlechter als die fehlerhafte Variante bewertet wurde, haben wir als fehlende Werte herausgenommen. Da wir diese fehlenden Werte nach unserer subjektiven Entscheidung und abhängig von einem anderen beobachteten Wert als fehlend interpretiert haben, klassifizierten wir sie als fehlende Werte, die zufällig

¹⁰ In der Skala für den Sicherheitsgrad wurden die Werte 1 als „sehr sicher“ und 4 als „sehr unsicher“ definiert.

zustande kommen bzw. als fehlende Werte der Art *Missing at Random* oder MAR.¹¹ Bei der Auswertung der Daten entschieden wir uns für das Ausschließen der fehlenden Werte bzw. für das Eliminierungsverfahren, was die Annahme eines MCAR oder MAR voraussetzt, und verwendeten die sogenannte *available-case* Methode.¹² Wir haben uns aus dem Grund für dieses Verfahren entschieden, weil wir die Nachteile des Imputationsverfahrens, nämlich die Unterschätzung der Varianz sowie der Streuung und die Manipulation der Daten durch Ersetzen eines nicht tatsächlich gegebenen Wertes, umgehen wollten (vgl. Hohl 2007:18). Bei der Darstellung der deskriptiven Daten in dieser Arbeit werden wir allerdings bei einigen Sätzen, die eine große Anzahl an fehlenden Werten aufweisen, überprüfen, ob zwischen den Ergebnissen, die durch das Ausschließen fehlender Werte zustande kommen, und den Resultaten, die durch das Ersetzen fehlender Werte durch den Mittelwert anderer vorhandener Werte geliefert werden, signifikante Unterschiede bestehen.

Bei der Eingabe von erkannten Fehlern wurden die irrelevanten Anmerkungen, z. B. Rechtschreibfehler, und bei der Eingabe des Sicherheitsgrads die Verweigerung der Antwort als fehlender Wert ausgeschlossen. Auch hier begnügten wir uns mit den vorhandenen Daten und bewerteten diese durch das Eliminierungsverfahren.

¹¹ In Bezug auf die Gründe für das Auftreten fehlender Werte werden drei Arten von Missing Data Mechanism definiert: *Missing Completely At Random* (MCAR), *Missing At Random* (MAR) und *Missing Not At Random* (MNAR). MCAR liegt vor, wenn die fehlenden Werte weder von einem beobachteten noch von einem unbeobachteten Wert abhängen. MAR bezeichnet fehlende Werte, die nur von einem tatsächlich beobachteten Wert abhängen. Und MNAR bezieht sich schließlich auf fehlende Werte, die von einem unbeobachteten Wert abhängig sind (vgl. Hohl 2007:14-17).

¹² Für die Behandlung fehlender Werte in der statistischen Auswertung gibt es zahlreiche Korrekturmethoden. Die zwei meist verwendeten Verfahren sind das Eliminierungsverfahren bzw. die Fallreduktion und das Ergänzungs- bzw. Imputationsverfahren. Die erste Methode, bei der der fehlende Wert aus der Analyse ausgeschlossen wird, unterteilt sich selbst in zwei Gruppen, nämlich die *Complete-Case*- und die *Available-Case-Analysis*. Bei der *Complete-Case* Methode werden nur Fälle mitberücksichtigt, die keine fehlenden Werte aufweisen. Diese Methode hat den Nachteil, dass dadurch viele Informationen verloren gehen. Bei der *Available-Case* Methode werden nur die einzelnen fehlenden Werte aus der Analyse herausgenommen und nicht die ganzen Fälle. Die Schwäche dieser Methode bezieht sich darauf, dass dadurch unterschiedliche Gruppengrößen entstehen, was allerdings bei einer geringen Anzahl von fehlenden Werten zu keinen Problemen bzw. Verzerrungen der Ergebnisse führen kann. Das zweite Verfahren, nämlich das Imputationsverfahren, wird angewendet, wenn die fehlenden Werte durch Ersatzwerte, z. B. durch den Mittelwert, den Median oder Modus anderer vorhandener Werte ergänzt werden (vgl. Messingschlager 2012:10-13).

Um die Daten unserer Fragebögen in den Codeplan für SPSS umzusetzen, haben wir darauf geachtet, dass die Wertelabels genau den Formulierungen des Fragebogens entsprechen. Wie im Anhang 3 ersichtlich ist, wurde im Beurteilungsbogen für jeden Satz auf die Frage „Wie gut wurde der Satz auf Standarddeutsch gesprochen?“ durch das Ankreuzen eines Wertes in folgender Skala geantwortet:

sehr gut 1 2 3 4 5 6 sehr schlecht

Da hier die Werte zwischen eins und sechs nicht verbalisiert dargestellt wurden, haben wir bei der Codeplanung auch nur die Werte 1 und 6 verbalisiert. Das Gleiche haben wir auch für die Angabe des Sicherheitsgrads bei der Fehlererkennung vorgenommen und nur die Werte 1 und 4 als sicher und unsicher in Worte gefasst.

Fehler: _____ sicher 1 2 3 4 unsicher

Die Antworten der Probanden auf die Frage „Welche Fehler haben Sie erkannt?“ wurden in SPSS in fünf Gruppen aufgeteilt. Da es sich hier um nominalskalierte Daten handelte, wurde die Antwortverweigerung bzw. die fehlende Antwort, was auch im Falle eines korrekten Satzes als fehlerfrei interpretiert wird, mit dem Wert 0 kodiert. Diesen Wert haben wir absichtlich nicht mit fehlerfrei gleichgesetzt, weil in vielen Fällen, in denen der Satz als sehr schlecht bzw. fehlerhaft bewertet wurde, keine Anmerkung zu den erkannten Fehlern gemacht wurde. Die richtige Erkennung des Fehlers wurde mit dem Wert 1 codiert und die Werte 2-5 bezogen sich auf die unterschiedlich falschen Erkennungen. Der Anhang 4 zeigt, wie unsere Variablen in dieser Arbeit codiert bzw. definiert wurden.

4. Auswertung der Ergebnisse und Überprüfung der Hypothesen

In den folgenden Kapiteln präsentieren wir die Ergebnisse der in dieser Arbeit durchgeführten Untersuchung zur Wirkung sprachlicher Fehler. Bevor wir mit der Prüfung unserer Hypothesen bzw. mit der inferenzstatistischen Analyse anfangen, präsentieren wir die deskriptiven Ergebnisse der Bewertung von 12 verschiedenen Fehlertypen, die jeweils aus sechs Sätzen bestanden und durch 32 Hörer beurteilt wurden. Durch diese Darstellung versuchen wir folgende Fragen zu beantworten:

- Wurden die Sätze in einer Fehlergruppe einheitlich bewertet oder gab es viele Ausreißer, die ein Gesamturteil über die Sätze erschweren?
- Wie groß ist der Unterschied zwischen den Bewertungen von Sätzen einer Fehlergruppe und den parallelen korrekten Sätzen?
- Wie häufig wurden Fehler in den jeweiligen Gruppen erkannt und wie sicher waren sich die Hörer bei dieser Erkennung?

Die erste Frage beantworten wir anhand der Präsentation von Maßen der zentralen Tendenz, d. h. Mittelwert, Median und Modalwert sowie Streuungsmaßen bzw. Standardabweichung. Da es sich bei unseren Daten um intervallskalierte Variablen handelt, überprüfen wir bei den einzelnen Gruppen, ob die Werte jeder Variable normalverteilt sind. Für die Überprüfung der Normalverteilung verwenden wir den Kolmogorov-Smirnov-Test¹³ (K-S-Test) und für die Visualisierung der Verteilungen präsentieren wir für jeden Satz Histogramme bzw. Polygonkurven, die auf dem beigefügten elektronischen Datenträger (CD-ROM) zu finden sind. Die Unterscheidung zwischen den normalen und nicht normalen Verteilungen sollte uns helfen, für die statistische Analyse die richtigen Testverfahren anzuwenden, denn für die Durchführung vieler statistischer Tests wird die Erfüllung des Kriteriums Normalverteilung vorausgesetzt.¹⁴

¹³ „Mit dem Kolmogorov-Smirnov-Test bei einer Stichprobe (Anpassungstest) wird die beobachtete kumulative Verteilungsfunktion für eine Variable mit einer festgelegten theoretischen Verteilung verglichen, die eine Normalverteilung, eine Gleichverteilung, eine Poisson-Verteilung oder Exponentialverteilung sein kann.“ (SPSS Base 16 Benutzerhandbuch. S. 424)

¹⁴ Wenn die Voraussetzung Normalverteilung nicht erfüllt ist, werden die sogenannten verteilungsfreien bzw. nichtparametrischen Testverfahren angewendet (vgl. Bortz & Schuster 2010:130).

Für die Beantwortung der zweiten Frage berechnen wir den Gesamtmittelwert¹⁵ jeder Fehlergruppe und der parallelen korrekten Sätze und überprüfen, ob sich diese Gruppen signifikant voneinander unterscheiden. Um den Unterschied zwischen den parallelen Sätzen besser zu beobachten, hätten wir auch einen Signifikanztest zwischen den einzelnen Sätzen durchführen können. Auf solch eine Analyse haben wir jedoch in dieser Arbeit verzichtet, da es sich bei der Fragestellung für die Hypothesen um den Unterschied zwischen den Fehlerarten und nicht die verschiedenen Ausprägungen eines Fehlers bzw. einzelner Sätze handelte.

Die Analyse der Differenzen zwischen den korrekten und den parallelen fehlerhaften Sätzen sollte uns helfen, die Hörbarkeit bzw. Auffälligkeit der Fehler überprüfen zu können. Eine Varianzanalyse für alle Fehlergruppen, welche wir später für die Überprüfung der Hypothesen durchführen, kann uns keinen solchen direkten Hinweis auf die Hörbarkeit der Fehler geben, denn eine schlechte Bewertung könnte z. B. nicht unbedingt auf die Auffälligkeit des Fehlers hindeuten, wenn der korrekte Satz genauso schlecht beurteilt wurde.

Die dritte Frage beantworten wir durch die Häufigkeitsanalyse der nominalskalierten Daten über die Anmerkungen der Probanden zu den von ihnen erkannten Fehlern und präsentieren diese durch die tabellarischen sowie graphischen Darstellungen. Bei der Präsentation des Sicherheitsgrads bei der Erkennung des Fehlers verzichten wir auf die Angaben, die bei einer falschen Erkennung gemacht wurden, und beschränken uns nur auf die Sicherheitsgrade bei korrekten Erkennungen.

¹⁵ Der Gesamtmittelwert jeder Fehlergruppe wurde in dieser Arbeit dadurch berechnet, dass zuerst für jede Person der Mittelwert ihrer Bewertungen für alle Sätze der jeweiligen Gruppe ausgerechnet wurde. Die neu berechneten Mittelwerte von 32 Personen bildeten eine neue Mittelwert-Variable, deren Mittelwert als Gesamtmittelwert der Fehlergruppe bezeichnet wurde.

4.1. Darstellung der Ergebnisse und Überprüfung der Hypothesen zu den phonologischen Fehlerarten

In diesem Kapitel werden zunächst die gewonnenen deskriptiven Daten zu den phonologischen Fehlerarten dargestellt. Bevor wir eine übersichtliche Darstellung der Unterschiede zwischen den sechs Aussprache-Fehlergruppen gewährleisten, versuchen wir, im nächsten Abschnitt durch eine Varianzanalyse herauszufinden, ob die Bewertungen dieser Fehlergruppen signifikant unterschiedlich sind.¹⁶ Um herauszufinden, wie die einzelnen Gruppen sich voneinander unterscheiden, führen wir anschließend eine Folgeanalyse durch und basierend auf deren Ergebnissen überprüfen wir die Verifizierung bzw. die Falsifizierung der Hypothesen 1.1, 1.2 und 1.3.

4.1.1. Deskriptive Werte für die Beurteilung der phonologischen Fehler

4.1.1.1. Fehlergruppe A: Realisierung der Affrikate [ts]

Die Wirkung der fehlerhaften Aussprache der Affrikate [ts] wurde anhand folgender Sätze, in denen das kursiv geschriebene Wort eine falsche Realisierung dieser Affrikate präsentierte, untersucht.

Satz 1.A. Ich habe leider keine *Zeit*.

Satz 2.A. Das *Justizlexikon* steht im Regal.

Satz 3.A. Er denkt nicht an seine *Zukunft*.

Satz 4.A. Meine Augen waren vom *Glanz* der Sonne geblendet.

Satz 5.A. Neulich ist *Benzin* viel teurer geworden.

Satz 6.A. Alle wissen, dass die Touristen sich mit Malaria *infziert* haben.

In allen Beispielen wurde die Affrikate [ts] fehlerhaft als Konsonant [s] ausgesprochen. Die folgende Tabelle präsentiert die wichtigsten statistischen Kennwerte für diese Sätze, die jeweils von 32 Probanden bewertet wurden.

¹⁶ Auf die Durchführung mehrerer Wilcoxon- bzw. t-Tests für die sechs Gruppen verzichten wir, da mit dem Anstieg der Testanzahl die Wahrscheinlichkeit, einen Fehler erster Art zu begehen, d. h. eine Hypothese fälschlicherweise zu verifizieren, ansteigt (vgl. Raab-Steiner & Benesch 2008:114).

(8) Statistische Kennwerte für die Fehlergruppe A

		Satz1.A	Satz2.A	Satz3.A	Satz4.A	Satz5.A	Satz6.A
N	gültig	31	28	30	26	28	31
	fehlend	1	4	2	6	4	1
Mittelwert		2,097	1,393	1,867	1,538	2,143	2,000
Median		2,000	1,000	2,000	1,000	2,000	2,000
Modalwert		2,0	1,0	1,0	1,0	1,0	2,0
Standardabweichung		,8309	,6853	,8604	,8115	1,0079	,8563

Wie der Tabelle (8) zu entnehmen ist, liegen die Durchschnittswerte zwischen 1,4 und 2,1 und die Abweichungen vom Mittelwert sind bei fast allen Sätzen kleiner als 1. Wenn wir unsere Skalenwerte in drei Gruppen aufteilen und die Werte 1 und 2 als gute Bewertung bezeichnen, können wir behaupten, dass die Fehler in der Gruppe A insgesamt positiv bzw. nicht störend bewertet wurden. Eine Übereinstimmung zwischen den drei Kennzahlen für die Lagemaße, d. h. zwischen dem Mittelwert, Median und Modalwert, ist nur beim Satz 6.A zu beobachten. Trotz dieser Übereinstimmung zeigte der K-S-Test auch für diesen Satz bzw. für alle Sätze Ergebnisse, die auf nicht normalverteilte Werte hindeuten¹⁷ (s. Anhang 5.A.1).

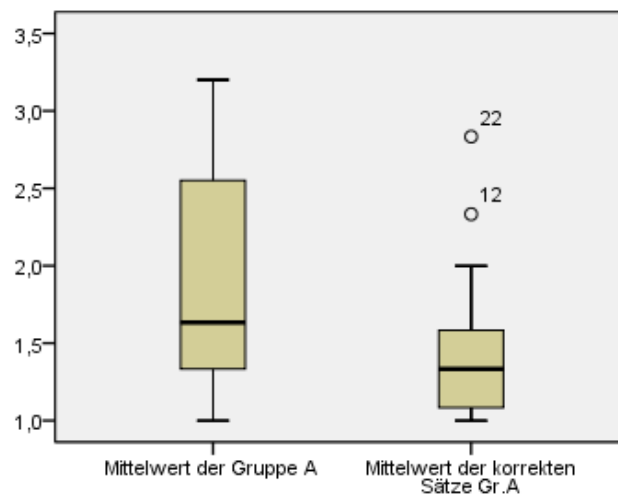
Für die Überprüfung der Hörbarkeit der Fehler dieser Gruppe haben wir die gleichen deskriptiven Kennwerte für die korrekten Sätze, d. h. für die Sätze 1 bis 6 berechnet (s. Anhang 6). Da uns an dieser Stelle nur der Gesamtunterschied zwischen der Gruppe A und der Gruppe mit korrekten Sätzen interessiert, verzichten wir auf eine Darstellung der deskriptiven Ergebnisse der einzelnen Sätze der Gruppe mit korrekten Sätzen und konzentrieren uns nur auf die Gesamtmittelwerte beider Gruppen. Bei unserer Analyse erwies die Gruppe A ($M = 1,86$, $SD = 0,65$) gegenüber der Gruppe mit korrekten Sätzen ($M = 1,43$, $SD = 0,43$) einen leicht höheren Mittelwert.

Da in manchen Fällen die nicht normalverteilten Variablen durch die Berechnung des Mittelwerts zu normalverteilten Variablen geändert werden, wurden die zwei Gesamtmittelwert-Variablen auch anhand der K-S-Tests auf Normalverteilung geprüft. Die Ergebnisse dieses Tests wiesen ebenso einen signifikanten Unterschied zu der Normalverteilung auf (s. Anhang 5.B). Aus diesem Grund wurde für die Überprüfung der Signifikanz des Unterschieds zwischen zwei Mittelwert-Variablen der nichtparametrische

¹⁷ Ein signifikantes Ergebnis beim K-S-Test weist darauf hin, dass die Daten nicht normalverteilt sind.

Test für zwei abhängige Stichproben, nämlich der Wilcoxon-Test¹⁸ angewendet.¹⁹ Auf Basis dieses Tests konnte ein signifikanter Unterschied zwischen diesen zwei Gruppen festgestellt werden ($z = -4,307$, $p = 0,000 < 0,01$, $r^{20} = -0,76^{21}$). Hier kann man den Schluss ziehen, dass Fehler bei der Realisierung der Affrikate [ts] mit einem signifikanten Unterschied zu den parallelen korrekten Beispielen gehört wurden. Das folgende Boxplot-Diagramm²² bestätigt diesen Unterschied.

(9) Vergleichende Darstellung des Medians und der Streuung für die Gruppe A und die Gruppe mit parallelen korrekten Sätzen



Die grafische Darstellung zeigt uns auf den ersten Blick, dass die rechte Box, die die

¹⁸ Beim Wilcoxon-Test werden anders als beim t-Test bei der Berechnung nicht das arithmetische Mittel, sondern die Rangplätze der Differenzen betrachtet (vgl. Raab-Steiner & Benesch 2008:125).

¹⁹ Für die Überprüfung der Signifikanz des Unterschieds zwischen zwei Mittelwert-Variablen verzichten wir in den folgenden Kapiteln auf die Darstellung der Ergebnisse des K-S-Tests, die im Anhang 5 zu finden sind, und verwenden für die nicht normalverteilten Variablen den Wilcoxon-Test und für die normalverteilten den t-Test.

²⁰ Da SPSS für nichtparametrische Tests keine Effektstärke berechnet, verwendeten wir für die Berechnung der Stärke der Differenzen zweier Gruppen, die mit dem Kleinbuchstaben r gekennzeichnet und mit Pearsons Korrelationskoeffizient identisch ist, die gleiche Formel wie beim Mann-Whitney-Test, nämlich $r = z / \sqrt{N}$. Hier symbolisiert N die Anzahl der Beobachtungen (vgl. Field 2013:234).

²¹ Für die Interpretation der Effektstärke beziehen wir uns auf die von J. Cohen (1988: 24–27, zit. nach Bortz & Schuster 2010:108) vorgeschlagenen Orientierungspunkte: $\delta = 0,2$ =kleiner Effekt, $\delta = 0,5$ =mittelgroßer Effekt und $\delta = 0,8$ =großer Effekt.

²² „Ein Boxplot ist eine Form der grafischen Darstellung von Daten. Es bildet die Verteilung einer einzelnen Variablen ab. Im Boxplot sind Median und Interquartilsabstand abgetragen. Es bietet eine gute Möglichkeit, die Rohdaten unverzerrt darzustellen und Ausreißer zu identifizieren.“ (Schäfer 2010:131)

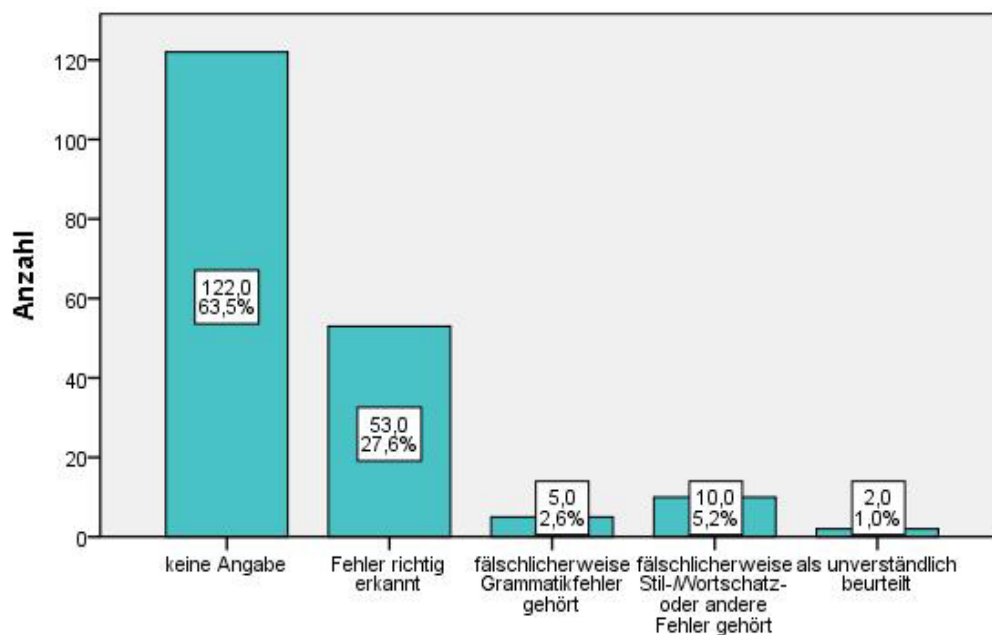
Verteilung für die Gruppe der korrekten Sätze präsentiert, kleiner ist als die linke Box für die Fehlergruppe A. Die größere Spannweite bzw. der größere Unterschied zwischen dem höchsten und dem niedrigsten Wert bei der Gruppe A weist auf eine höhere Streuung der Daten dieser Gruppe hin. Außerdem ist der Zentralwert bzw. Median, der durch einen Strich in der Mitte der Box gekennzeichnet ist, bei der rechten Gruppe niedriger.

Die Tatsache, dass die berechneten Mittelwerte und Standardabweichungen sich nur wenig voneinander unterscheiden und die oben grafisch dargestellten Boxen einen größeren Unterschied zwischen den beiden Gruppen zeigen, kann nur auf die Existenz der Ausreißer in der Gruppe mit korrekten Sätzen, die sich außerhalb der Box befinden, zurückgeführt werden. Der Mittelwert und die Standardabweichung der Gruppe mit korrekten Sätzen zeigen im Vergleich zu dem dargestellten Median und der Streuung im Boxplot einen höheren Wert, weil sie bei der Berechnung auch die Ausreißer mitberücksichtigt haben. Hier stellt sich die Frage, wie man die Lage- und Streuungsmaße der Daten am besten präsentieren kann. Da die Berechnung der deskriptiven Daten „Mittelwert“ und „Standardabweichung“ für die weiteren Analysen unabdingbar ist, wäre es von Vorteil, als komplementäre Information auch die Boxplots zu präsentieren. Dadurch können wir erkennen, ob wir in unseren Daten Ausreißer haben, die die deskriptiven Daten beeinflussen können.

Für die Darstellung der Angaben der Probanden zu den erkannten Fehlern begnügen wir uns hier mit der Veranschaulichung der Ergebnisse der Gesamtgruppe A, die wir durch das Zusammenfassen aller Anmerkungen von sechs Sätzen dieser Gruppe in eine Variable²³ analysiert haben. Das folgende Balkendiagramm präsentiert die absolute sowie relative Häufigkeit der fünf verschiedenen Antwortausprägungen für die Gruppe A.

²³ Da wir die Werte von sechs Sätzen, die jeweils von 32 Probanden bewertet wurden, in eine Variable zusammengefasst haben, bestand die neue Variable aus 192 Fällen und so wurde in der neuen Variable $n=192$.

(10) Fehlerangaben der Probanden für die Gruppe A



Wie das Diagramm zeigt, haben bei der Fehlergruppe A nur 27,6% der Probanden den Fehler richtig erkannt, wobei von dieser Gruppe 45,3 % ihre Erkennung als ganz sicher bezeichnet haben (s. Anhang 7). Die folgende Tabelle präsentiert die Häufigkeit der korrekten Erkennung für die einzelnen Sätze.

(11) Absolute und relative Häufigkeit der korrekten Erkennung des Fehlers für die Sätze der Gruppe A

	Satz1.A	Satz2.A	Satz3.A	Satz4.A	Satz5.A	Satz6.A
Fehler richtig erkannt	18	4	12	0	6	13
	56,3%	12,5%	37,5%	0	18,8%	40,6%

Wie der Tabelle (11) zu entnehmen ist, erhielten die Sätze 1.A, 6.A und 3.A die höchsten Prozentzahlen der richtigen Erkennung. Die falsche Aussprache von [ts] im Wort *Glanz* im Satz 4.A. wurde von keinem Probanden erkannt und nur 12% der Probanden erkannten die fehlerhafte Aussprache von [ts] im Wort *Justizlexikon*. Die Unauffälligkeit der falschen Realisierung der Affrikate [ts] in diesen zwei Wörtern könnte damit begründet werden, dass in beiden Wörtern die Affrikate am Wort- bzw. Silbenende auftraten. Inwieweit diese Begründung tatsächlich auf die Position der Affrikate im Wort zurückzuführen ist, bedarf meiner Meinung nach einer detaillierten Untersuchung, deren Durchführung über den Rahmen unserer Arbeit hinausgehen würde.

4.1.1.2. Fehlergruppe B: Auslautverhärtung

Um die fehlerhafte Aussprache der stimmhaften Frikative am Ende der Silbe zu präsentieren, verwendeten wir die kursiv geschriebenen Wörter in folgenden Sätzen.

Satz 7.B. Meine Freundin ist *naiv*, aber nicht dumm. (/f/→/v/)

Satz 8.B. Der Wind *blies* mir ins Gesicht. (/s/→/z/)

Satz 9.B. Er beteiligt sich immer *aktiv* am Unterricht. (/f/→/v/)

Satz 10.B. Einer der Hunde war sehr *aggressiv* und hat andere Hunde verletzt. (/f/→/v/)

Satz 11.B. Leider *wies* die Rechnung einige Fehler auf. (/s/→/z/)

Satz 12.B. Wenn du dich mehr informieren willst, *lies* dieses Buch. (/s/→/z/)

Wie in den Klammern dargestellt ist, wurden bei diesen Beispielen die Frikative am Wortende anstatt als stimmlose [f] und [s] jeweils stimmhaft bzw. als [v] und [z] ausgesprochen. Die folgende Tabelle präsentiert die deskriptiven Ergebnisse dieser Sätze.

(12) Statistische Kennwerte für die Fehlergruppe B

	Satz7.B	Satz8.B	Satz9.B	Satz10.B	Satz11.B	Satz12.B
N gültig	29	30	14	27	27	27
fehlend	3	2	18	5	5	5
Mittelwert	1,276	1,367	1,429	2,000	1,778	1,407
Median	1,000	1,000	1,000	2,000	2,000	1,000
Modalwert	1,0	1,0	1,0	1,0	1,0	1,0
Standardabweichung	,5276	,6149	,6462	1,1435	,8916	,6939

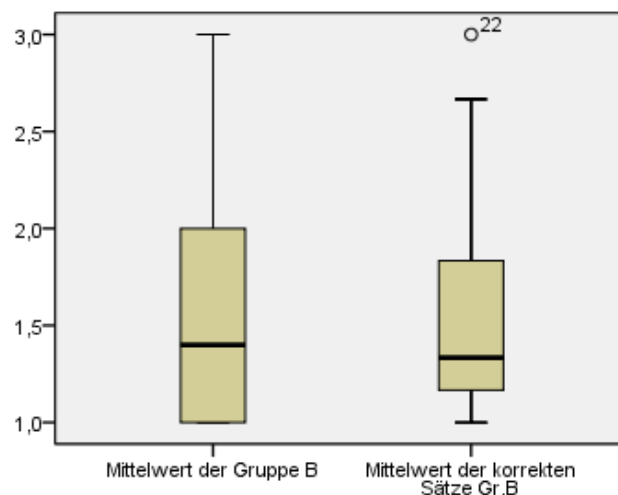
Wie die Tabelle zeigt, liegen die Mittelwerte in dieser Gruppe zwischen 1,2 und 2 und die Standardabweichung ist bei fast allen Sätzen kleiner als 1. Die Ergebnisse des K-S-Tests sowie die asymmetrische Form der Polygonkurven weisen bei allen Sätzen eine Verletzung der Normalverteilung auf (s. Anhang 5.A.1). Auch hier könnte man im Allgemeinen behaupten, dass die Fehler dieser Gruppe als relativ positiv bzw. nicht störend bewertet wurden.

Auffällig an dieser Tabelle ist die große Anzahl der fehlenden Werte für den Satz 9.B. Bei diesem Satz haben 18 Personen die korrekte Version des Satzes schlechter bewertet als die fehlerhafte, weshalb ihre Bewertungen aus der Analyse herausgenommen wurden. Da bei diesem Satz ca. 56% der Werte fehlen, haben wir auch die Kennzahlen für die Variante mit

ersetzten Werten analysiert. Um zu überprüfen, ob es zwischen beiden Varianten bzw. der Variante mit fehlenden Werten und der mit vollständigen Daten einen signifikanten Unterschied gibt, führten wir einen Wilcoxon-Test durch. Das Ergebnis dieses Tests ($z = 0,000$, $p = 1,000 > 0,05$, $r = 0$) zeigte keinen signifikanten Unterschied zwischen beiden Versionen. Aus diesem Grund und um die Einheitlichkeit der Analyse beizubehalten, haben wir die erste Variante, nämlich die Variante mit fehlenden Werten in unserer weiteren Analyse mitberücksichtigt.

Die statistische Analyse der Gruppe B ergab einen Gesamtmittelwert von ($M = 1,60$, $SD = 0,64$), der im Vergleich zum Gesamtmittelwert der parallelen korrekten Sätze ($M = 1,55$, $SD = 0,55$) einen sehr geringen Unterschied aufzeigte. Die Mittelwert-Variablen wurden auch hier anhand eines Wilcoxon-Tests miteinander verglichen. Das Ergebnis dieses Tests ($z = -1,183$, $p = 0,237 > 0,05$, $r = -0,20$) bestätigte die Annahme, dass die Fehler in der Gruppe B gegenüber der Gruppe mit korrekten Sätzen kaum gehört bzw. als störend bewertet wurden. Das folgende Boxplot-Diagramm veranschaulicht diese Annahme.

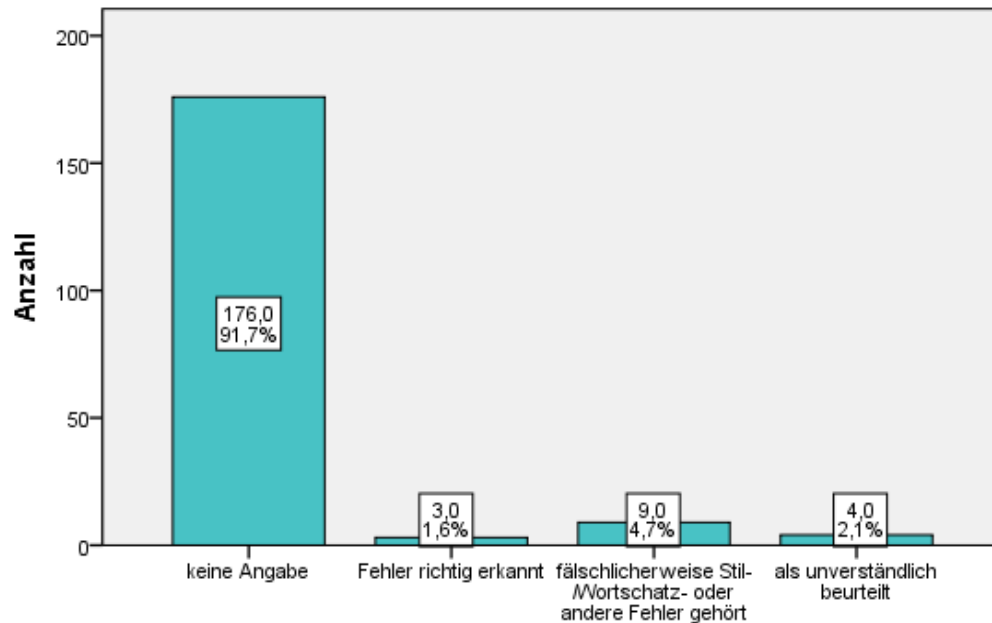
(13) Vergleichende Darstellung des Medians und der Streuung für die Gruppe B und die Gruppe mit parallelen korrekten Sätzen



Obwohl die Gruppe B im Vergleich zu der Gruppe mit korrekten Sätzen eine größere Streuung aufweist, kann man an diesem Diagramm sehen, dass der Median in beiden Gruppen fast auf der gleichen Linie liegt. Die Unauffälligkeit der Auslautverhärtungsfehler wurde auch bei der Analyse der Anmerkungen der Probanden zu den erkannten Fehlern merkbar. Laut dieser Analyse haben nur 1,6 % der Probanden die fehlerhafte Aussprache der Frikative am Wortende als Fehler erkannt, wobei niemand seine korrekte Erkennung als ganz sichere Anmerkung bzw. mit dem Sicherheitsgrad 1 bestätigte (s. Anhang 7). Nachfolgend

präsentieren wir die Häufigkeitsverteilung der Anmerkungen zu der Gruppe B sowie die Häufigkeit der korrekten Erkennungen bei den einzelnen Sätzen.

(14) Fehlerangaben der Probanden für die Gruppe B



(15) Absolute und relative Häufigkeit der korrekten Erkennung des Fehlers für die Sätze der Gruppe B

	Satz7.B	Satz8.B	Satz9.B	Satz10.B	Satz11.B	Satz12.B
Fehler richtig erkannt	0	1	0	1	1	0
	0	3,1%	0	3,1%	3,1%	0

4.1.1.3. Fehlergruppe C: Vokalquantität

Die falsche Aussprache der Vokallänge wurde anhand der kursiv geschriebenen Wörter in folgenden Sätzen dargestellt.

13. C. Wie lange kannst du deinen *Atem* anhalten? (/a:/ → /a/)

14. C. Sie ist die gute *Seele* in unserem Team. (/e:/ → /ε/)

15. C. Mein *Zahn* ist sehr empfindlich und tut beim Kauen weh. (/a:/ → /a/)

16. C. Es wird *ewig* dauern, bis er seine Schulden bezahlt hat. (/e:/ → /ε/)

17.C. Am Wochenende sollte das *Wetter* schön werden. (/ε / → /ε:/)

18.C. Die Kinder spielen mit dem *Ball*, obwohl sie es gar nicht dürfen. (/a/ → /a:/)

Die in Klammern dargestellten Anmerkungen zeigen, wie die unterstrichenen Vokale unterschiedlich lang ausgesprochen wurden. Die folgende Tabelle präsentiert die deskriptiven Ergebnisse der Bewertungen dieser Sätze.

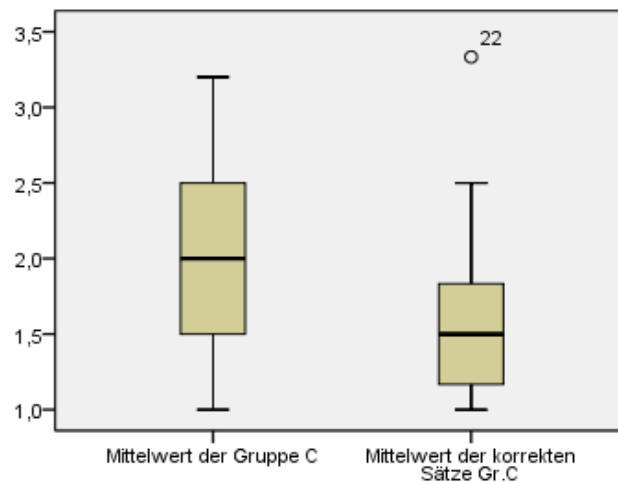
(16) Statistische Kennwerte für die Fehlergruppe C

		Satz13.C	Satz14.C	Satz15.C	Satz16.C	Satz17.C	Satz18.C
N	gültig	29	24	24	29	29	32
	fehlend	3	8	8	3	3	0
Mittelwert		2,000	1,500	2,542	2,241	2,414	1,625
Median		2,000	1,000	2,000	2,000	2,000	1,500
Modalwert		2,0	1,0	2,0	2,0	2,0	1,0
Standardabweichung		1,0351	,5898	1,0206	,9124	1,0183	,7071

Der Mittelwertbereich zwischen 1,5 und 2,5 weist auch hier auf eine relativ gute Bewertung bzw. nicht störende Wirkung der Fehler in dieser Gruppe hin. Wie der Tabelle (16) zu entnehmen ist, gibt es bei fast allen Sätzen eine relative Übereinstimmung zwischen den drei Maßzahlen der zentralen Tendenz und die Streuung der Bewertungen des Mittelwerts beträgt bei allen Sätzen maximal 1,03. Trotz dieser Übereinstimmung zeigte der K-S-Test für alle Sätze p-Werte, die auf nicht normalverteilte Daten hinwiesen (s. Anhang 5.A.1).

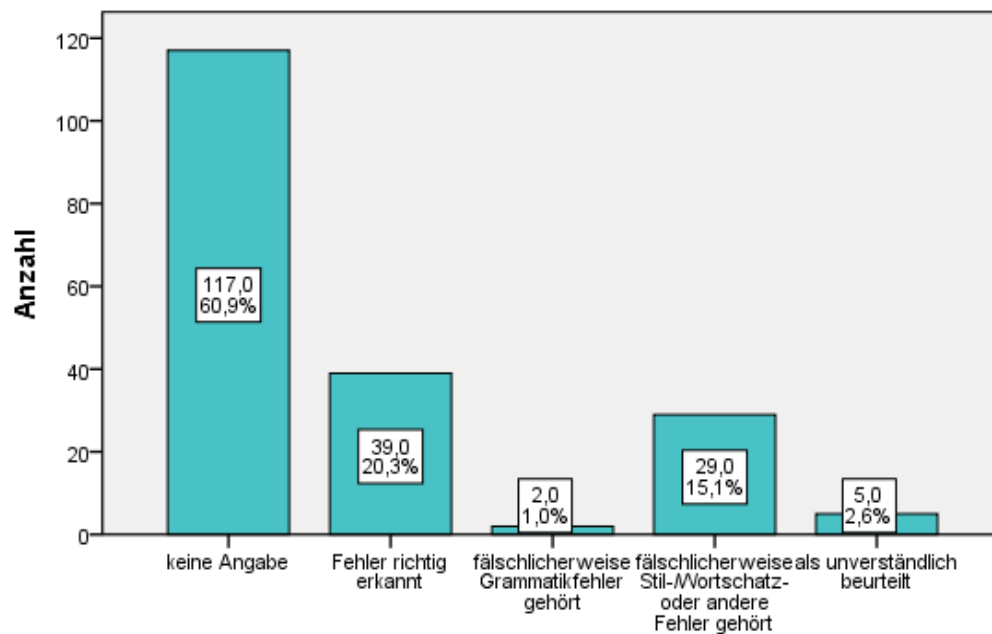
Für die Gruppe C wurde ein Gesamtmittelwert von ($M = 2.06$, $SD = 0,67$) berechnet. Die Analyse der parallelen korrekten Sätze zu dieser Gruppe ergab einen Gesamtmittelwert von ($M = 1.61$, $SD = 0,54$). Laut den Ergebnissen des Wilcoxon-Tests ($z = -4,191$, $p = 0,000 < 0,01$, $r = -0,74$) konnte ein signifikanter Unterschied zwischen diesen Gruppen festgestellt werden. Die visualisierte Darstellung des Medians und der Streuung durch folgendes Boxplot-Diagramm bestätigt diese Annahme.

(17) Vergleichende Darstellung des Medians und der Streuung für die Gruppe C und die Gruppe mit parallelen korrekten Sätzen



Im folgenden Balkendiagramm können wir zudem sehen, wie die relativen sowie absoluten Häufigkeiten der Anmerkungen der Hörer zu den erkannten Fehlern in sechs Sätzen der Gruppe C aussehen.

(18) Fehlerangaben der Probanden für die Gruppe C



Wie wir sehen, wurden nur 20,3% der Vokalquantität-Fehler, davon 48,7% mit dem Sicherheitsgrad 1, richtig erkannt (s. Anhang 7). Wenn wir die Analyse der einzelnen Sätze in Bezug auf Fehlererkennung in der folgenden Tabelle betrachten, stellt sich heraus, dass die meisten richtigen Erkennungen für die Sätze 17.C (ca. 65%) und 18.C (ca. 28%) registriert wurden. Damit scheint der fehlerhafte Ersatz eines Kurzvokals durch einen Langvokal für

deutsche Muttersprachler auffälliger bzw. einfacher zu erkennen zu sein. Diese Vermutung bildet eine eigene Hypothese, deren Überprüfung über den Rahmen dieser Arbeit hinausgeht.

(19) Absolute und relative Häufigkeit der korrekten Erkennung des Fehlers für die Sätze der Gruppe C

	Satz13.C	Satz14.C	Satz15.C	Satz16.C	Satz17.C	Satz18.C
Fehler richtig	0	1	4	4	21	9
erkannt	0	3,1%	12,5%	12,5%	65,6%	28,1%

Die Tatsache, dass wir bei der Bewertung dieser Sätze, wie die Tabelle (16) zeigt, keinen so großen Unterschied zwischen den deskriptiven Kennzahlen feststellen konnten, könnte damit begründet werden, dass ein Fehler, wenn er richtig erkannt wurde, nicht immer auch unbedingt negativ bewertet wird. Diese Annahme bedarf allerdings einer ausführlichen Analyse, die wir bei der Prüfung der Hypothese 5 durchführen werden.

4.1.1.4. Fehlergruppe D: Realisierung des Umlauts

Für die falsche Realisierung des Umlauts haben wir die in den folgenden Sätzen markierten Wörter angewendet.

19. D. Die Gastgeber servierten ein üppiges Mahl. (/Y/→/o/)

20. D. Kein Problem. Das nehme ich dir nicht übel. (/y:/→ /u:/)

21. D. Sie sind sehr gütig zu mir. (/y:/→ /u:/)

22. D. Sie hat sich gar nicht bemüht. (/y:/→ /u:/)

23. D. Jeden Tag treffen sich die Künstler. (/Y/→/o/)

24. D. Auch wenn es hier sehr gemütlich ist, müssen wir bald wieder aufbrechen. (/y:/→ /u:/)

In diesen Beispielen wurden die Umlaute [y:] und [Y] jeweils ersetzt durch die Vokale [u:] und [o]. Die folgende Übersicht zeigt die deskriptiven Werte für sechs Sätze dieser Gruppe.

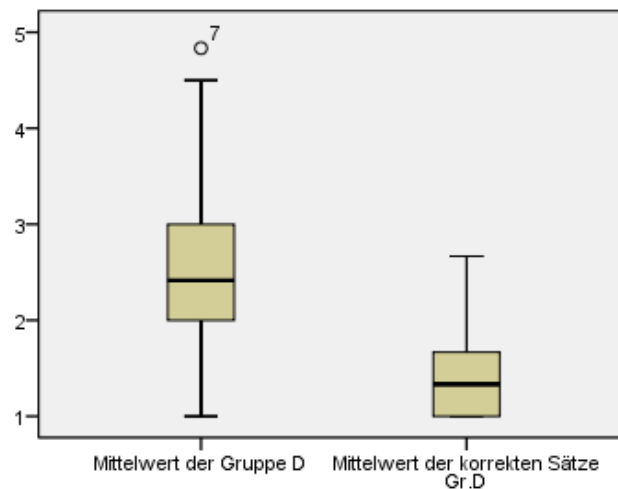
(20) Statistische Kennwerte für die Fehlergruppe D

		Satz19.D	Satz20.D	Satz21.D	Satz22.D	Satz23.D	Satz24.D
N	gültig	32	32	32	32	31	31
	fehlend	0	0	0	0	1	1
Mittelwert		2,531	2,469	2,750	2,594	2,613	2,419
Median		2,000	2,000	3,000	2,000	2,000	2,000
Modalwert		2,0	2,0	2,0	2,0	2,0	2,0
Standardabweichung		1,0468	,9499	1,0776	,9108	1,0223	,8860

Wie die Tabelle zeigt, erhielten die fehlerhaften Sätze dieser Gruppe im Vergleich zu den anderen bis jetzt ausgewerteten Fehlergruppen höhere Mittelwerte. Der kleine Unterschied zwischen den Mittelwerten der einzelnen Sätze sowie die relative Übereinstimmung der Maßzahlen der zentralen Tendenz bei jedem Satz weisen auf eine einheitliche Bewertung der gehörten Sätze dieser Gruppe hin. Obwohl die Polygonkurven aller Sätze in dieser Gruppe eine glockenförmige Kurve, die als Hinweis auf die Normalverteilung der Daten gelten sollte, aufzeigte, zeigte der K-S-Test für alle Sätze Ergebnisse, die von der Normalverteilung abwichen (s. Anhang 5.A.1).

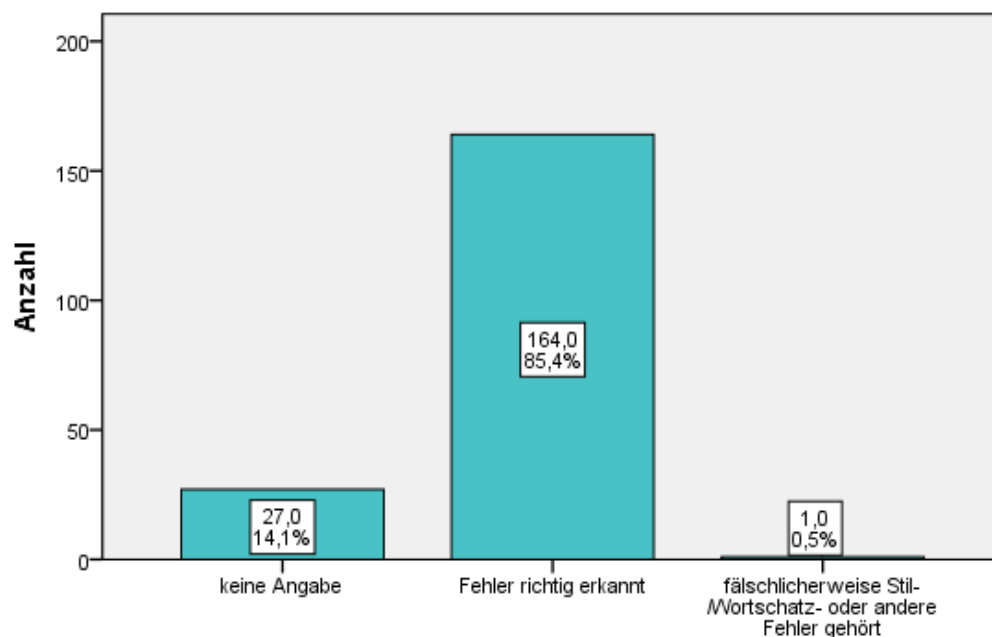
Die Berechnung des Gesamtmittelwerts der Gruppe D ergab einen Wert von ($M = 2.55$, $SD = 0,84$), der im Vergleich zu der Gruppe der parallelen korrekten Sätze ($M = 1.44$, $SD = 0,46$) einen höheren Wert aufwies. Um zu überprüfen, ob dieser Unterschied bedeutsam bzw. signifikant ist, wurde wiederum ein Wilcoxon-Test durchgeführt. Das Ergebnis dieses Tests ($z = -4,86$, $p = 0,000 < 0,01$, $r = -0,86$) deutete auf einen signifikanten Unterschied zwischen den zwei Gruppen. Diese deutliche Ungleichheit ist auch im folgenden Boxplot-Diagramm leicht zu erkennen.

(21) Vergleichende Darstellung des Medians und der Streuung für die Gruppe D und die Gruppe mit parallelen korrekten Sätzen



Die Analyse der Anmerkungen der Probanden zu den Fehlern dieser Gruppe ergab, dass 85,4% der Fehler in der Realisierung des Umlauts richtig erkannt wurden. Das folgende Balkendiagramm veranschaulicht dieses Ergebnis.

(22) Fehlerangaben der Probanden für die Gruppe D



Wie das Diagramm zeigt, wurden die Fehler in der Gruppe D im Vergleich zu den anderen bis jetzt analysierten Fehlergruppen besser erkannt. Die korrekte Erkennung von 85,4% wurde hier in 73,8% der Fälle mit dem Sicherheitsgrad 1 bewertet (s. Anhang 7). Wenn wir die folgende Tabelle betrachten, können wir sehen, dass die Prozentzahlen der richtigen Erkennung der Fehler in den einzelnen Sätzen der Gruppe D relativ einheitlich sind.

(23) Absolute und relative Häufigkeit der korrekten Erkennung des Fehlers für die Sätze der Gruppe D

	Satz19.D	Satz20.D	Satz21.D	Satz22.D	Satz23.D	Satz24.D
Fehler richtig erkannt	23 71,9%	28 87,5%	29 90,6%	29 90,6%	29 90,6%	26 81,3%

4.1.1.5. Fehlergruppe E: Konsonantenhäufung in der Wortanlautsilbe

Die abweichende Aussprache der Wörter, die in ihrer Wortanlautsilbe einen Konsonantencluster aufweisen, wurde mithilfe der in den folgenden Sätzen markierten Wörter präsentiert.

- 25. E. Deutsch ist für mich eine zweite *Sprache*.
- 26. E. Wir sind auf der *Straße*.
- 27. E. Ich habe Angst vor *Schlangen*.
- 28. E. Die Verbrennungen haben ihr unerträgliche *Schmerzen* bereitet.
- 29. E. Letztes Jahr habe ich mein *Studium* beendet.
- 30. E. Obwohl er nicht so groß ist, kann er drei Meter hoch *springen*.

Diese Wörter, die hauptsächlich aus mit [ʃ] anlautenden Konsonantengruppen bestanden, wurden durch das Voranstellen des Sprossvokals [e]²⁴ vor die ganze Konsonantengruppe verändert ausgesprochen. Die folgende Tabelle veranschaulicht die deskriptiven Kennzahlen für diese Gruppe.

(24) Statistische Kennwerte für die Fehlergruppe E

	Satz25.E	Satz26.E	Satz27.E	Satz28.E	Satz29.E	Satz30.E
N gültig	28	31	31	30	31	31
fehlend	4	1	1	2	1	1
Mittelwert	2,964	3,000	2,161	1,400	2,226	2,645
Median	3,000	3,000	2,000	1,000	2,000	3,000
Modalwert	3,0	2,0	2,0	1,0	2,0	2,0
Standardabweichung	1,1701	1,0954	,7788	,7240	1,2030	1,0181

²⁴ Der eingeschobene Vokal wurde bei einigen Beispielen auch als [ɛ] und [ə] ausgesprochen.

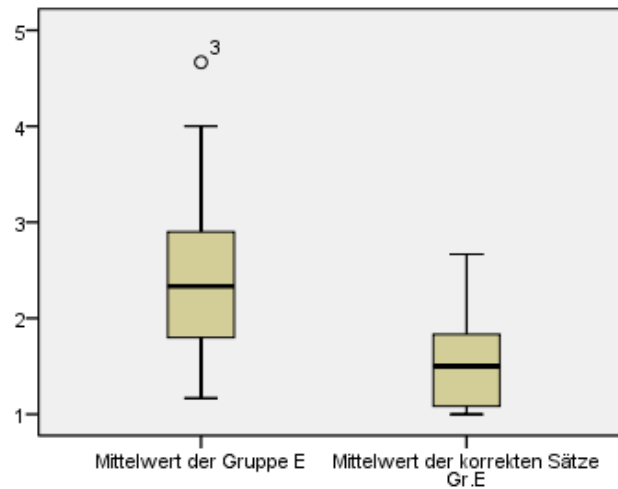
Wie der Übersicht zu entnehmen ist, erhielten die abweichenden Realisierungen der Wörter *Straße* im Satz 26.E, *Sprache* im Satz 25.E und *springen* im Satz 30.E die höchsten Mittelwertzahlen. Die negative Bewertung dieser Wörter kann auf die größere Anzahl der Konsonanten in ihrer Anlautsilbe zurückgeführt werden. Auch in der im Rahmen meiner Masterarbeit durchgeführten Untersuchung wies das Wort *Straße* im Vergleich zu anderen Wörtern eine größere Prozentzahl der als abweichend beurteilten Aussprache auf. Diese Annahme kann die Vermutung anregen, dass persische Deutschsprecher die Wörter, die in ihrer Anlautsilbe eine große Anzahl Konsonanten aufweisen, besonders abweichend aussprechen.

Wenn wir die Kennzahlen der zentralen Tendenz dieser Sätze betrachten, können wir feststellen, dass diese Maßzahlen abgesehen von den Sätzen 26.E und 30.E bei allen anderen Sätzen eine Übereinstimmung aufweisen. Obwohl bei einigen Sätzen wie bei 25.E, 26.E und 30.E eine glockenförmige Kurve zu sehen war, zeigten alle Sätze dieser Gruppe laut K-S-Test Ergebnisse, die eine Abweichung von der Normalverteilung aufwiesen (s. Anhang 5.A.1).

Die Gruppe E zeigte mit dem Gesamtmittelwert von ($M = 2,38$, $SD = 0,80$) im Vergleich zu der Gruppe der parallelen korrekten Sätze ($M = 1,53$, $SD = 0,46$) einen höheren Mittelwert auf. Die Mittelwert-Variablen wurden hier als normalverteilt klassifiziert, weshalb wir einen parametrischen Test, nämlich den t-Test, angewendet haben. Laut den Ergebnissen dieses Tests wurden die abweichend gesprochenen Sätze deutlich schlechter beurteilt als die parallelen korrekten Sätze ($t(31) = 8,20$, $p = 0,000 < 0,01$, $r^{25} = 0,82$). Das folgende Boxplot-Diagramm bestätigt diese Annahme.

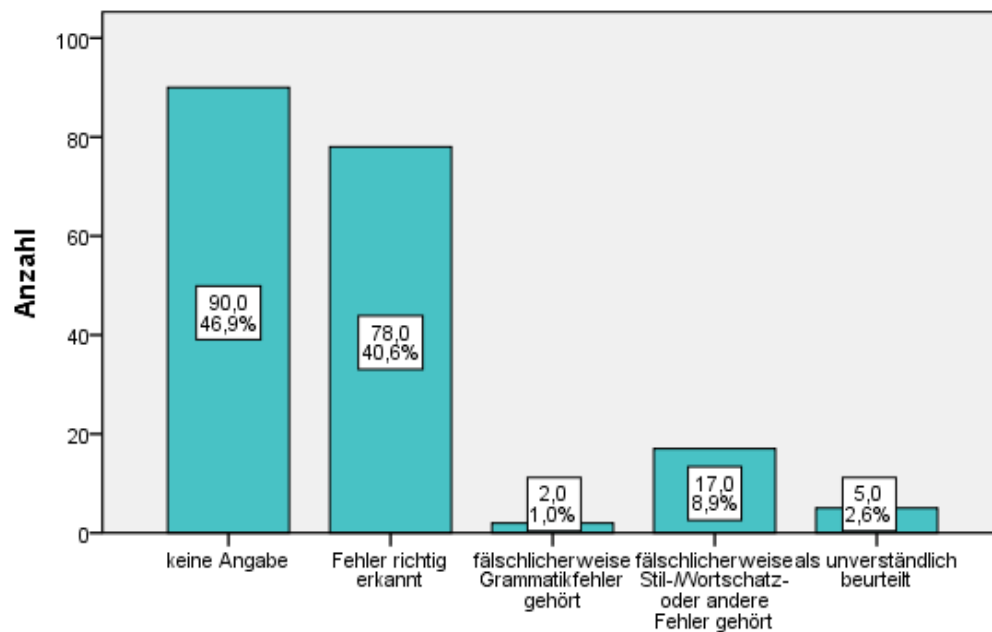
²⁵ SPSS berechnet auch für den t-Test mit abhängigen Stichproben keine Effektstärke, weshalb wir folgende Formel verwendet haben: $r = \sqrt{t^2 / (t^2 + df)}$ (vgl. Field 2013:386).

(25) Vergleichende Darstellung des Medians und der Streuung für die Gruppe E und die Gruppe mit parallelen korrekten Sätzen



Die Häufigkeitsanalyse der Anmerkungen der Probanden zu den erkannten Fehlern der Gruppe E, die wir anhand des folgenden Balkendiagramms präsentieren, wies eine Prozentzahl von 40,6% der richtigen Erkennung auf.

(26) Fehlerangaben der Probanden für die Gruppe E



Die korrekten Erkennungen wurden nur in 41% der Fälle als sicher (Skalenwert 1) angekreuzt (s. Anhang 7). Obwohl die Fehler in dieser Gruppe ähnlich wie in der Gruppe D als störend und auffällig bewertet wurden, sehen wir hier, dass eine große Anzahl der Probanden (ca.

47%) auf eine Anmerkung zu dem erkannten Fehler verzichtet haben. Die richtige Erkennung bei den einzelnen Sätzen lässt sich anhand folgender Tabelle darstellen.

(27) Absolute und relative Häufigkeit der korrekten Erkennung des Fehlers für die Sätze der Gruppe E

	Satz25.E	Satz26.E	Satz27.E	Satz28.E	Satz29.E	Satz30.E
Fehler richtig erkannt	15 46,9%	21 65,6%	11 34,4%	1 3,1%	13 40,6%	17 53,1%

Auch hier wiesen die Wörter *Straße* im Satz 26.E, *springen* im Satz 30.E und *Sprache* im Satz 25.E die höchste Anzahl der korrekt erkannten Fehler auf. Diese Übereinstimmung spricht, im Gegensatz zu den Ergebnissen der Gruppe C, für einen positiven Zusammenhang zwischen der korrekten Erkennung des Fehlers und dessen negativer Bewertung, was wir bei der Analyse der Hypothese 5 ausführlich besprechen werden.

4.1.1.6. Fehlergruppe F: Wortakzent

Die in den vorliegenden Sätzen markierten Komposita wurden als Stimuli für die Überprüfung der Wirkung falscher Wortakzentpositionen angewendet.

31. F. Wir treffen uns am *Hauptbahnhof*.
32. F. Die Kinder sind auf dem *Spielplatz*.
33. F. Meine Freundin ist immer noch im *Krankenhaus*.
34. F. Der Taxifahrer hat an der *Bushaltestelle* gehalten.
35. F. Als ausländischer Student muss man in Deutschland zuerst einen *Sprachkurs* besuchen.
36. F. Ich wusste nicht, dass das *Studienkolleg* hier in der Nähe ist.

In diesen Beispielen wurden die fett markierten Wortteile bzw. bei zweiteiligen Komposita die zweite und bei dreigliedrigen Zusammensetzungen die mittlere Konstituente inkorrekt akzentuiert. Die folgende Tabelle präsentiert uns, wie die einzelnen Sätze insgesamt bewertet wurden.

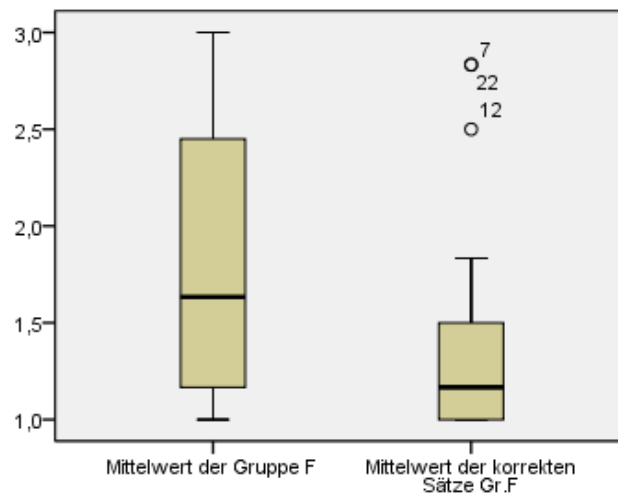
(28) Statistische Kennwerte für die Fehlergruppe F

		Satz31.F	Satz32.F	Satz33.F	Satz34.F	Satz35.F	Satz36.F
N	gültig	25	29	32	31	28	31
	fehlend	7	3	0	1	4	1
Mittelwert		1,480	1,759	1,813	1,806	1,929	1,806
Median		1,000	2,000	1,500	2,000	2,000	2,000
Modalwert		1,0	1,0	1,0	1,0	1,0	1,0
Standardabweichung		,7141	,8305	1,0607	,8725	,8133	,8334

Wie aus der Tabelle zu entnehmen ist, besteht bei keinem Satz eine Übereinstimmung zwischen den drei Kennwerten der zentralen Verteilungslage, nämlich zwischen dem Mittelwert, Median und Modalwert. Das liegt daran, dass die Verteilungen in diesen Sätzen, wie es auch vom K-S-Test bestätigt wurde (s. Anhang 5.A.1), nicht-symmetrisch bzw. nicht normalverteilt sind. Wenn wir die Annahme, dass bei unsymmetrischen Verteilungen der Median aufgrund seiner geringen Anfälligkeit gegenüber Ausreißern die Lage der Verteilung besser wiedergibt, berücksichtigen, können wir feststellen, dass hier der Wert 2 als Mitte der Verteilung am häufigsten vertreten ist (vgl. Schäfer 2010). Das Streuungsmaß der Sätze weist keine großen Unterschiede auf, weshalb wir von einer relativ einheitlichen Bewertung der Probanden ausgehen können.

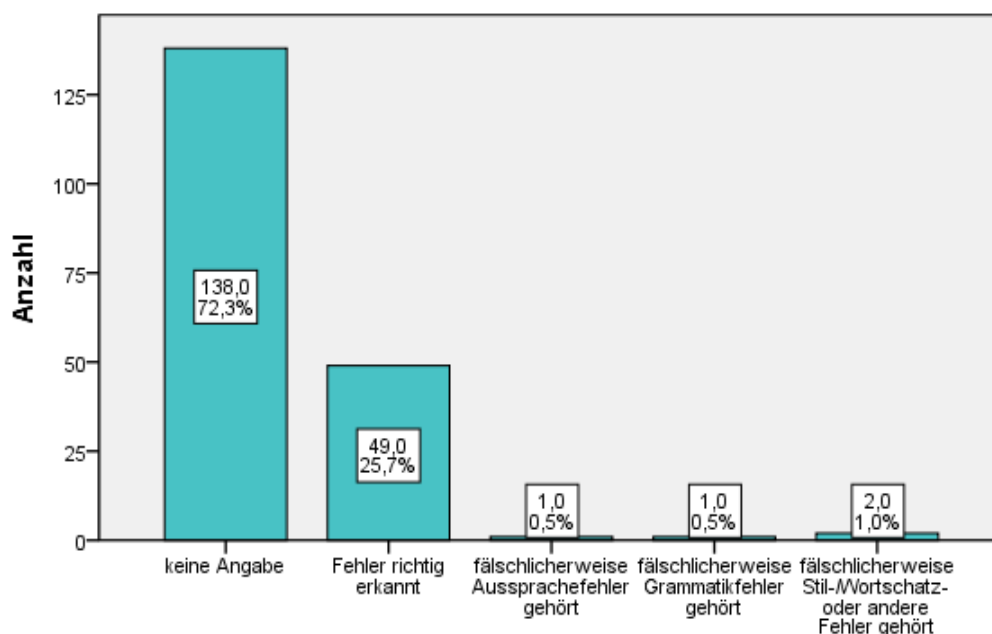
Die sechs Sätze der Gruppe F ergaben einen Gesamtmittelwert von ($M = 1,78$, $SD = 0,69$), der im Vergleich zu der parallelen Gruppe der korrekten Sätze ($M = 1,36$, $SD = 0,50$) leicht höher war. Das Ergebnis des Wilcoxon-Tests ($z = -3,80$, $p = 0,000 < 0,01$, $r = -0,67$) zeigte auf, dass die Wortakzentfehler mit einem signifikanten Unterschied zu den parallelen korrekten Beispielen gehört wurden. Das folgende Boxplot-Diagramm bestätigt dieses Ergebnis.

(29) Vergleichende Darstellung des Medians und der Streuung für die Gruppe F und die Gruppe mit parallelen korrekten Sätzen



Hier kann man aufgrund der Unempfindlichkeit des Medians gegenüber Ausreißern, die bei der Gruppe der korrekten Sätze außerhalb der Box zu sehen sind, den Unterschied zwischen diesen zwei Gruppen deutlicher erkennen. Wie wir sehen, weist die Gruppe der korrekten Sätze im Vergleich zu der Gruppe F einen niedrigeren Median und eine geringere Streuung auf. Die Auswertung der Anmerkungen der Probanden zu den erkannten Fehlern ergab, dass insgesamt nur in 25,7% der Fälle, davon nur 33% mit dem Sicherheitsgrad 1 (s. Anhang 7), die falsche Akzentposition tatsächlich als Wortakzentfehler angemerkt wurde.

(30) Fehlerangaben der Probanden für die Gruppe F



Die folgende Tabelle präsentiert uns, wie die richtige Erkennung bei den einzelnen Sätzen aussah.

(31) Absolute und relative Häufigkeit der korrekten Erkennung des Fehlers für die Sätze der Gruppe F

	Satz31.F	Satz32.F	Satz33.F	Satz34.F	Satz35.F	Satz36.F
Fehler richtig erkannt	4 12,5%	9 28,1%	9 28,1%	11 34,4%	14 43,8%	2 6,3%

Warum die falsche Akzentuierung der Wörter *Studienkolleg* im Satz 36.F und *Sprachkurs* im Satz 35.F so unterschiedlich erkannt wurde, bedarf meiner Meinung nach einer ausführlichen Untersuchung, deren Durchführung den Rahmen dieser Arbeit sprengen würde.

4.1.2. Statistische Auswertung der gewonnenen Daten über die Bewertung der phonologischen Fehlerarten

4.1.2.1. Varianzanalyse und die Rangordnung der Aussprache-Fehlergruppen

Da unsere Mittelwert-Variablen außer den Fehlergruppen C und E von der Normalverteilung abweichen, müssen wir Verfahren auswählen, die gegenüber dieser Abweichung unempfindlich sind. Die parametrischen Signifikanzverfahren der ANOVA (Analysis of Variance) setzen eine Normalverteilung der Messwerte in allen untersuchten Gruppen voraus. Obwohl in vielen Statistik-Büchern die Varianzanalyse als robust gegenüber der Verletzung dieser Voraussetzung gilt, haben wir uns in dieser Arbeit für ein nichtparametrisches Verfahren entschieden (vgl. Schäfer 2011:135).

Diese Entscheidung begründete sich mit der Annahme, dass in unserer Untersuchung die unabhängigen Variablen latent²⁶ bzw. verborgen sind. Die Durchführung einer parametrischen ANOVA-Analyse bedarf der Eingabe der unabhängigen Variablen, was bei nichtparametrischen Tests nicht notwendig ist. Aus diesem Grund und da es bei unseren

²⁶ Nach dem Gesichtspunkt der Messbarkeit lassen sich Variablen in zwei Gruppen, nämlich manifeste und latente Variablen unterteilen. Manifeste Variablen können direkt gemessen werden und manifestieren sich als registrierte Zahlen bzw. Daten, z. B. das Alter. Latente Variablen sind dagegen nach außen nicht sichtbar und liegen im Verborgenen, z. B. verschiedene Fehlertypen in unserer Arbeit. Obwohl die Fehlerarten unsere unabhängigen Variablen sind, die die abhängigen Variablen, nämlich die Bewertung der Probanden, beeinflussen, können wir sie nicht direkt als eine Variable in der Analyse mitberücksichtigen, weil wir sie nicht messen konnten (vgl. Schäfer 2010:29).

Daten um den Vergleich von mehr als zwei abhängigen Stichproben handelte, wendeten wir den Friedman-Test²⁷ an. Die folgende Tabelle präsentiert die Ranginformationen der sechs untersuchten Gruppen, wobei sich laut den Ergebnissen dieses Tests ($\chi^2(5)^{28} = 85,85$, $p = 0,000 < 0,01$) mindestens zwei Gruppen signifikant voneinander unterscheiden.

(32) Friedman-Test: Ranginformation zu den phonologischen Fehlergruppen

	Mittlerer Rang
Mittelwert der Gruppe A	2,86
Mittelwert der Gruppe B	1,81
Mittelwert der Gruppe C	3,88
Mittelwert der Gruppe D	5,20
Mittelwert der Gruppe E	4,75
Mittelwert der Gruppe F	2,50

Der Friedman-Test berechnet die mittleren Rangplätze dadurch, dass er zuerst für die Bewertungen jeder Person für die sechs Gruppen, die in unserer Arbeit als Mittelwert der Bewertungen der sechs Sätze manifestiert wurden, Rangplätze vergibt. Die mittleren Ränge ergeben sich durch das Addieren der Rangplätze pro Spalte bzw. pro Gruppe dividiert durch die Anzahl der Personen (vgl. Raab-Steiner & Benesch 2008:127).

Da in unserer Arbeit die Anzahl der Skalenwerte und der Ränge beide 6 bzw. identisch sind, können wir die oben angeführten mittleren Ränge als Mittelwert der Bewertungen für jede Gruppe betrachten. Diese Annahme erlaubt uns, die Ergebnisse des Friedman-Tests für mittlere Ränge als eine Basis-Information für die Klassifikation unserer Fehlerarten hinsichtlich ihrer störenden Wirkung auf die deutschen Muttersprachler zu verwenden. Wenn wir diese Ränge der Größe nach ordnen und sie durchnummerieren, ergibt sich die folgende Rangliste.

²⁷ Der Friedman-Test ist ein nichtparametrisches Äquivalent einer Varianzanalyse mit Messwiederholungen. Dieser Test überprüft die Rangplätze der Differenzen und nicht wie beim ANOVA-Test das arithmetische Mittel der untersuchten Gruppen (vgl. Raab-Steiner & Benesch 2008:125).

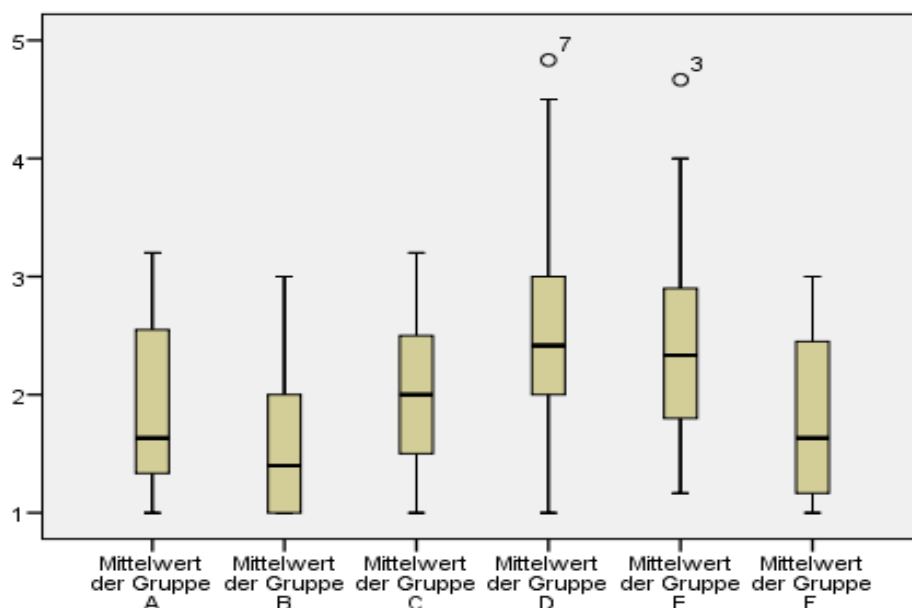
²⁸ Der Freiheitsgrad für den Friedman-Test wird durch die Formel $(n-1)$ berechnet, wobei n die Anzahl der Gruppen bezeichnet (vgl. Field 2013:255).

(33) Rangordnung der Aussprachefehler hinsichtlich ihrer Wirkung auf die deutschen Muttersprachler

Platz	Gruppe	Mittlerer Rang
1	Gruppe D (Realisierung des Umlauts)	5,20
2	Gruppe E (Konsonantenhäufung in der Wortanlautsilbe)	4,75
3	Gruppe C (Vokalquantität)	3,88
4	Gruppe A (Realisierung der Affrikate [ts])	2,86
5	Gruppe F (Wortakzent)	2,50
6	Gruppe B (Auslautverhärtung)	1,81

Die höheren Zahlen für mittlere Ränge weisen hier, wie bei unseren Skalenwerten, eine schlechtere Bewertung auf. Dies bedeutet, dass die Fehler der Gruppe D bzw. die falsche Realisierung des Umlauts in unserer Untersuchung von den deutschen Hörern am wenigsten toleriert wurden. Wenn wir die Boxplot-Diagramme bzw. die Median-Veranschaulichung dieser Gruppen nebeneinander betrachten, können wir die gleichen Rangplätze für die sechs Aussprache-Fehlergruppen feststellen. Die vergleichende Übersicht der Mediane und Mittelwerte dieser sechs Gruppen ist im Anhang 8 zu finden.

(34) Vergleichende Darstellung des Medians und der Streuung für die phonologischen Fehlerarten



Um zu überprüfen, ob unsere Fehlergruppen hinsichtlich ihrer Differenzen zu den parallelen korrekten Sätzen den gleichen Rangplatz belegen, haben wir für die Variablen, die aus den Differenzen der Mittelwerte der Fehlergruppen und der entsprechenden korrekten Sätze

bestanden,²⁹ den Friedman-Test durchgeführt. Auch bei den Differenz-Variablen lag der p-Wert unter dem Signifikanzniveau von 0,01 ($\chi^2(5) = 86,83$, $p = 0,000$), so dass wir davon ausgehen können, dass sich mindestens zwei Gruppen signifikant voneinander unterscheiden. Die Tabelle der Rangplatz-Informationen wies eine ähnliche Rangreihe wie die Tabelle (32) auf.

(35) Friedman-Test: Ranginformationen zu den Differenzen der phonologischen Fehlergruppen und parallelen korrekten Sätze

	Mittlerer Rang
Differenz A-K ³⁰ A	2,98
Differenz B-KB	1,66
Differenz C-KC	3,27
Differenz D-KD	5,34
Differenz E-KE	4,81
Differenz F-KF	2,94

Wie aus der Tabelle ersichtlich ist, zeigte auch hier die Gruppe D die höchste Anzahl der Differenz zwischen ihrer Fehlergruppe und der entsprechenden Gruppe der korrekten Sätze, was darauf hindeutet, dass diese Gruppe mehr als andere Fehlergruppen als fehlerhaft empfunden wurde. Wenn wir die Rangplätze der Größe nach gliedern, ergibt sich eine ähnliche Rangfolge wie in der Tabelle (33), nämlich die Gruppe D gefolgt von den Gruppen E, C, A, F und B. Betrachtet man nun die Anzahl der korrekten Erkennung der Fehler und die dazu angegebenen Sicherheitsgrade, zeigen sich einige Unterschiede.

(36) Relative Häufigkeit der korrekten Erkennung der Aussprachefehler und die als ganz sicher angekreuzte Prozentzahl der korrekten Erkennung

Gruppen	A	B	C	D	E	F
Prozentzahl der korrekten Erkennung	27,6%	1,6%	20,3%	85,4%	40,6%	25,7%
Prozentzahl der mit Sicherheitsgrad 1 angekreuzten korrekten Erkennung	45,3%	0	48,7%	73,8%	41%	33%

²⁹ Um dies zu berechnen, wurde für jede Person der Mittelwert ihrer Bewertungen für die sechs fehlerhaften Sätze vom Mittelwert ihrer Bewertungen für die sechs parallelen korrekten Sätze subtrahiert. Dadurch entstanden für die Gruppen A-F sechs neue Variablen, die als Differenz-Variablen bezeichnet wurden.

³⁰ Der Buchstabe K symbolisiert bei allen Gruppen die Gruppe der korrekten Sätze.

Ähnlichkeiten zwischen der Anzahl der korrekten Erkennung und der erstellten Reihenfolge in der Tabelle (33) betreffen vor allem die Gruppe D, gefolgt von E mit der höchsten und der Gruppe B mit der niedrigsten Prozentzahl der korrekten Erkennung. Bei den Angaben zum Sicherheitsgrad weisen nur zwei Gruppen, nämlich die Gruppen D und B, mit jeweils der höchsten und niedrigsten Prozentzahl den gleichen Rangplatz auf.

Diese Unterschiede führen auf die Vermutung zurück, dass ein Fehler, wenn er richtig erkannt wird, nicht immer auch negativ bewertet werden würde. Die Untersuchung des Zusammenhangs zwischen der korrekten Erkennung und der negativen Bewertung vollziehen wir später bei der Überprüfung der Hypothese 5. Wovon wir an dieser Stelle definitiv ausgehen können, ist die Tatsache, dass von den deutschen Muttersprachlern laut den Ergebnissen dieser Studie die Fehler bei der Realisierung des Umlauts und der Konsonantenhäufung in der Wortanlautsilbe weniger toleriert werden als die anderen Aussprache-Fehlerarten und die Verletzung der Regel Auslautverhärtung von den meisten Hörern kaum als Fehler wahrgenommen wird.

4.1.2.2. Paarvergleiche bzw. Post-hoc-Analyse der Aussprache-Fehlergruppen

Der signifikante p-Wert des Friedman-Tests im vorigen Kapitel zeigte uns, dass mindestens zwei Aussprache-Fehlergruppen signifikant verschieden sind. Zur Beantwortung der Frage, welche Fehlergruppen sich tatsächlich voneinander unterscheiden und ob unsere Rangfolge der Fehlergruppen auf signifikanten Unterschieden basiert, führen wir nun eine weitere Analyse bzw. Post-hoc-Analyse³¹ durch. Die Tabelle (37) stellt die Ergebnisse der einzelnen Paarvergleiche dieser Analyse dar.

³¹ Das in dieser Arbeit verwendete Post-hoc-Verfahren wird in SPSS durch eine weitere Funktion der nichtparametrischen Tests für verbundene Stichproben namens „paarweise Vergleiche“ durchgeführt und vergleicht die einzelnen Paare durch einen Vorzeichentest (vgl. Field 2013:256).

(37) Paarweiser Vergleich der Bewertung phonologischer Fehlerarten

Stichprobe1-Stichprobe2	Test-statistik	Std. Fehler	Standard Test-statistik	Sig.	Angep. Sig.
Mittelwert der Gruppe B-Mittelwert der Gruppe F	-,688	,468	-1,470	,142	1,000
Mittelwert der Gruppe B-Mittelwert der Gruppe A	1,047	,468	2,238	,025	,378
Mittelwert der Gruppe B-Mittelwert der Gruppe C	-2,062	,468	-4,410	,000	,000
Mittelwert der Gruppe B-Mittelwert der Gruppe E	-2,938	,468	-6,281	,000	,000
Mittelwert der Gruppe B-Mittelwert der Gruppe D	-3,391	,468	-7,249	,000	,000
Mittelwert der Gruppe F-Mittelwert der Gruppe A	,359	,468	,768	,442	1,000
Mittelwert der Gruppe F-Mittelwert der Gruppe C	1,375	,468	2,940	,003	,049
Mittelwert der Gruppe F-Mittelwert der Gruppe E	2,250	,468	4,811	,000	,000
Mittelwert der Gruppe F-Mittelwert der Gruppe D	2,703	,468	5,780	,000	,000
Mittelwert der Gruppe A-Mittelwert der Gruppe C	-1,016	,468	-2,171	,030	,448
Mittelwert der Gruppe A-Mittelwert der Gruppe E	-1,891	,468	-4,042	,000	,001
Mittelwert der Gruppe A-Mittelwert der Gruppe D	-2,344	,468	-5,011	,000	,000
Mittelwert der Gruppe C-Mittelwert der Gruppe E	-,875	,468	-1,871	,061	,921
Mittelwert der Gruppe C-Mittelwert der Gruppe D	-1,328	,468	-2,840	,005	,068
Mittelwert der Gruppe E-Mittelwert der Gruppe D	,453	,468	,969	,333	1,000

Wie wir sehen, gibt uns SPSS verschiedene Informationen über die einzelnen Vergleiche, wobei wir uns für die Feststellung der signifikanten Unterschiede auf die Informationen in der letzten Spalte bzw. die angepassten Signifikanzniveaus beziehen müssen. Laut den Ergebnissen dieser Tabelle unterscheiden sich nur die acht Fehlergruppen-Paare, die markiert sind, signifikant voneinander. Wenn wir unsere Rangordnung für die Fehlergruppen in der Tabelle (33), nämlich die Reihenfolge D, E, C, A, F und B, betrachten, können wir feststellen, dass die Fehlergruppen, die untereinander angeordnet sind, sowie die Gruppen, die relativ nah zueinander platziert sind, nämlich die Paare D&C und A&B, keine signifikanten Unterschiede aufweisen. Schauen wir auf die signifikanten p-Werte, können wir die folgenden Schlussfolgerungen ziehen:

1. Die Fehlergruppen D und E wurden signifikant schlechter beurteilt als die Gruppen A, B und F. 2. Die Gruppe C wurde signifikant schlechter beurteilt als die Gruppen F und B.

Wenn wir das Ergebnis der Post-hoc-Analyse der Differenz-Variablen im Anhang 9 betrachten, sehen wir einige minimale Unterschiede. Die Differenz-Variable C zeigte bei dieser Analyse keinen signifikanten Unterschied zu der Differenzgruppe F und unterschied sich signifikant von den Gruppen E und D.

Da die Fehlergruppe C sich in beiden Post-hoc-Analysen unterschiedlich verhält, beschränken wir die Schlussfolgerungen der ersten Analyse auf das erste Ergebnis und gehen definitiv davon aus, dass die Gruppen mit Fehlern in der Realisierung des Umlauts (Gruppe D) und der Konsonantenhäufung (Gruppe E) signifikant schlechter beurteilt wurden als die Gruppen mit der falschen Akzentposition (Gruppe F), der fehlerhaften Aussprache der Affrikate [ts] (Gruppe A) und der Verletzung der Auslautverhärtungsregel (Gruppe B).

4.1.3. Überprüfung der Hypothesen zu den phonologischen Fehlern

Die bisherige Vorauswertung der Bewertungen von Aussprache-Fehlergruppen sollte uns ermöglichen, die in den Hypothesen 1.1, 1.2 und 1.3 formulierten Vermutungen, die sich auf den Unterschied zwischen zwei Gruppen beziehen, zu überprüfen. Für die Untersuchung dieser Hypothesen greifen wir auf ein passendes Prüfverfahren zurück, das für den Vergleich zweier abhängiger Stichproben geeignet ist. Die Ergebnisse der im letzten Kapitel durchgeführten Post-hoc-Analyse können wir nur für die Überprüfung der Hypothese 1.3 verwenden, da es sich nur bei dieser Hypothese um einen direkten Vergleich zweier Fehlergruppen, nämlich der Gruppen C und D, handelt.

4.1.3.1. Hypothese 1.1

In der Hypothese 1.1 wurde vermutet, dass die suprasegmentalen Fehler von deutschen Muttersprachlern schlechter beurteilt werden als die segmentalen. In der Gruppe der suprasegmentalen Fehler bezogen wir in dieser Arbeit die Gruppen mit Fehlern in der Realisierung der Vokalquantität (Gruppe C), der Konsonantenhäufung in der Wortanlautsilbe (Gruppe E) und des Wortakzents (Gruppe F) ein. Die Gruppe der segmentalen Fehler bestand aus den Gruppen mit Fehlern in der Produktion der Affrikate [ts] (Gruppe A), der stimmhaften Frikative am Silbenende (Gruppe B) und des Umlauts (Gruppe D). Wenn wir

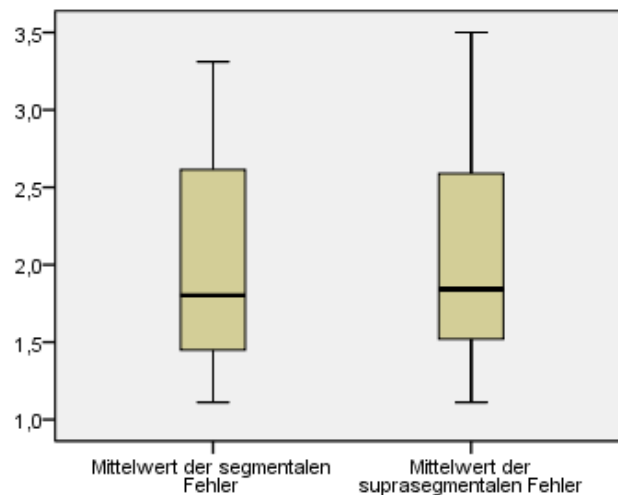
diese Aufteilung in eine statistische Notation umschreiben, lautet unsere Hypothese: $M (C+E+F) > M (A+B+D)$, wobei die höhere Anzahl des Mittelwerts auf eine schlechtere Bewertung hinweist.

Für die Untersuchung dieser Hypothese bildeten wir zwei neue Variablen. Die erste Variable präsentierte die Bewertung der suprasegmentalen Fehlergruppen und bestand aus Mittelwerten der Bewertungen der Gruppen C, E und F. Die zweite Variable wurde aus Mittelwerten der Bewertungen der Gruppen A, B und D, d. h. der Gruppen der segmentalen Fehler gebildet.

Obwohl die Gruppe mit suprasegmentalen Fehlern einen leicht höheren Mittelwert ($M = 2,07$, $SD = 0,67$) im Vergleich zu der Gruppe mit segmentalen Fehlern ($M = 2,00$, $SD = 0,66$) aufzeigte, wies das Ergebnis des Wilcoxon-Tests ($z = -1,86$, $p = 0,063 > 0,05$, $r = -0,33$) keinen signifikanten Unterschied zwischen diesen zwei Gruppen auf. Wenn wir das festgelegte Signifikanzniveau von 0,05 als Maßstab für die Ablehnung oder Bestätigung der Hypothesen annehmen, müssen wir davon ausgehen, dass die Hypothese 1.1 als nicht bestätigt gilt. Beziehen wir uns jedoch auf die Annahme, dass die Festlegung des Signifikanzniveaus eine subjektive Entscheidung des Forschers ist und in einigen Studien gelegentlich auch ein p-Wert von 0,1 als Signifikanzniveau angenommen werden kann (vgl. Bortz & Schuster 2010:101), können wir unseren p-Wert von 0,06, der sich in der Nähe des konventionellen Signifikanzniveaus von 0,05 befindet, als signifikant betrachten.

Um den Lageunterschied zwischen diesen zwei Gruppen besser zu beobachten, schauen wir uns auch die Mediane bzw. das Boxplot-Diagramm der beiden Variablen an. Wie aus dem folgenden Diagramm ersichtlich ist, unterscheiden sich die Medianpositionen beider Gruppen nur ganz minimal voneinander.

(38) Vergleichende Darstellung des Medians und der Streuung für die Mittelwert-Variablen der segmentalen und suprasegmentalen Fehlerarten



Da die Bewertungsdifferenz zwischen den fehlerhaften und den korrekten Sätzen für die Überprüfung der Unterschiede zwischen den Fehlergruppen eine nicht zu unterschätzende Rolle spielt, verglichen wir auch die parallelen Differenz-Variablen miteinander. Für die Differenz-Variable der suprasegmentalen Fehlergruppe wurde ein Mittelwert von ($M = 0,57$, $SD = 0,40$) berechnet, der nur ganz leicht über dem Mittelwert der Differenz-Variable der segmentalen Fehlergruppe ($M = 0,53$, $SD = 0,34$) lag. Der Wilcoxon-Test berechnete für den Vergleich beider Mittelwerte einen p-Wert ($z = -0,916$, $p = 0,36 > 0,05$, $r = -0,16$), der deutlich über dem Signifikanzniveau lag, weswegen die Feststellung, dass es zwischen den segmentalen und den suprasegmentalen Fehlergruppen keinen Unterschied gibt, noch einmal bestätigt wurde.

Wenn wir auf die Ergebnisse der paarweisen Vergleiche der Aussprache-Fehlergruppen, die in der Tabelle (37) präsentiert wurden, zurückgreifen und nur die signifikanten Ergebnisse berücksichtigen, können wir feststellen, dass nur die vier folgenden Vergleiche sich auf den signifikanten Unterschied zwischen den segmentalen und den suprasegmentalen Fehlergruppen beziehen: Vergleich der Mittelwerte der Gruppen (B & C), (B & E), (A & E) und (F & D). Betrachten wir die festgestellte Reihenfolge der Wichtigkeit der Fehlerarten hinsichtlich ihrer Bewertung durch die deutschen Muttersprachler im vorigen Kapitel, können wir feststellen, dass nur beim Vergleich zwischen den Gruppen F & D die segmentale Fehlerart, d. h. die Gruppe D, signifikant schlechter als die suprasegmentale Gegengruppe bewertet wurde.

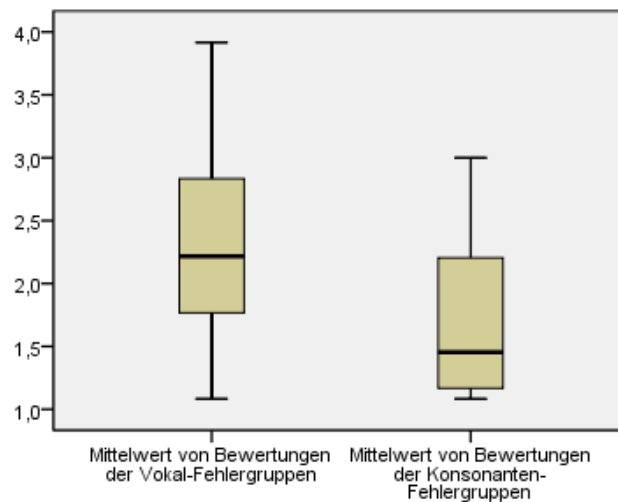
In allen anderen Vergleichen wurden die suprasegmentalen Gruppen signifikant schlechter bewertet als die segmentalen. Da die Gruppe D unter den Aussprache-Fehlergruppen die höchste Mittelwertanzahl erreicht hat und die einzige Gruppe der segmentalen Fehlergruppen ist, die signifikant schlechter als die suprasegmentalen Gruppen bewertet wurde, liegt an dieser Stelle die Vermutung nahe, dass der Mittelwert der Bewertungen der segmentalen Fehlergruppe definitiv viel niedriger gewesen wäre, hätten wir diese Fehlergruppe aus der Analyse herausgenommen, und so hätte diese Gruppe signifikant mit besseren Bewertungen abschneiden können als die suprasegmentale Fehlergruppe.

4.1.3.2. Hypothese 1.2

Bei der Formulierung der Hypothese 1.2 sind wir davon ausgegangen, dass die Aussprachefehler in der Realisierung der Vokale von Muttersprachlern als störender wahrgenommen werden als die fehlerhafte Produktion der Konsonanten. Für die Vokalfehlergruppe bildeten wir eine Variable aus den Mittelwerten der Bewertungen der Gruppen mit Vokalquantitätsfehlern (Gruppe C) und mit abweichender Aussprache des Umlauts (Gruppe D). Die Konsonantenfehlergruppe war eine Variable aus den Mittelwerten der Bewertungen der Gruppen mit falscher Realisierung der Affrikate [ts] (Gruppe A) und mit Verletzung der Auslautverhärtungsregel (Gruppe B). Die statistische Umschreibung dieser Hypothese lautete von daher: $M(C+D) > M(A+B)$.

In unserer Analyse zeigte die Vokalfehlergruppe einen relativ höheren Mittelwert ($M = 2,31$, $SD = 0,70$) im Vergleich zu der Fehlergruppe der Konsonanten ($M = 1,73$, $SD = 0,62$). Auch hier wurde für die Überprüfung der Signifikanz dieses Unterschieds ein Wilcoxon-Test durchgeführt. Laut den Ergebnissen dieses Tests ($z = -4,90$, $p = 0,000 < 0,01$, $r = -0,86$) wurde die Gruppe mit Vokalfehlern hochsignifikant schlechter beurteilt als die Konsonanten-Fehlergruppe. Dieser Unterschied ist auch durch die Visualisierung der Lage des Medians im folgenden Boxplot-Diagramm gut zu erkennen.

(39) Vergleichende Darstellung des Medians und der Streuung für die Mittelwert-Variablen der vokalischen und konsonantischen Fehlergruppen



Um zu überprüfen, ob die Differenz-Variablen der Vokal- sowie Konsonantenfehlergruppen sich genauso signifikant voneinander unterscheiden, vollzogen wir das gleiche Testverfahren für die Differenz-Variablen. Diese Analyse ergab für die vokalische Differenz-Variable einen Mittelwert von ($M = 0,78$, $SD = 0,41$), der im Vergleich zu der Differenz-Variable der Konsonantenfehlergruppe ($M = 0,24$, $SD = 0,27$) relativ höher lag. Auch hier wurde anhand eines t-Tests ein hochsignifikanter Unterschied ($t(31) = -9,66$, $p = 0,000 < 0,01$, $r = 0,86$) zwischen den Differenz-Variablen der Gruppen (A+B) und (C+D) nachgewiesen. Laut diesem Ergebnis können wir davon ausgehen, dass die Hypothese 1.2 definitiv als bestätigt gelten kann.

4.1.3.3. Hypothese 1.3

In der Hypothese 1.3 wurde angenommen, dass die fehlerhafte Aussprache der Vokallänge (Gruppe C) von deutschen Muttersprachlern negativer bewertet wird als die falsche Realisierung der Vokalqualität (Gruppe D). Wie auch aus der festgelegten Reihenfolge der Fehlerarten hinsichtlich ihrer Wichtigkeit für die deutschen Hörer im Kapitel 4.1.2.1. ersichtlich ist, wies die Gruppe D einen höheren Mittelwert der Bewertungen auf ($M = 2,55$, $SD = 0,84$) als die Gruppe C ($M = 2,06$, $SD = 0,67$). Laut diesem Ergebnis wurde die Gruppe mit Fehlern in der Realisierung der Vokallänge (Gruppe C) sogar besser beurteilt als die Gruppe mit Fehlern in der Vokalqualität (Gruppe D). Die bessere Bewertung der Gruppe C wird auch anhand der Mediane beider Gruppen in der Abbildung (34) erkennbar.

Um zu sehen, ob die Gruppe C signifikant besser bewertet wurde, betrachten wir zuerst die Ergebnisse der Post-hoc-Analyse in der Tabelle (37). Bei diesem paarweisen Vergleich wurde der Unterschied zwischen diesen zwei Gruppen mit einem Signifikanzniveau von $p = 0,068$ als nicht signifikant beurteilt. Führen wir jedoch einen Wilcoxon-Test durch, zeigt sich dieser Unterschied als hochsignifikant ($z = -3,893$ $p = 0,000 < 0,01$, $r = -0,68$).

Die gegensätzlichen Signifikanzniveaus in beiden Testverfahren lassen sich nur damit begründen, dass bei der Post-hoc-Analyse das Signifikanzniveau aufgrund des paarweisen Mehrfachvergleichs an die Anzahl der Vergleiche angepasst wird. Da durch diese Anpassung die Wahrscheinlichkeit, einen Fehler erster Art zu begehen, d. h. eine Hypothese fälschlicherweise zu verifizieren, verringert wird, beziehen wir uns hier auf das Ergebnis der Post-hoc-Analyse und gehen davon aus, dass die Gruppe C mit einem sich dem konventionellen Signifikanzniveau annähernden P-Wert von 0,68 nicht signifikant besser bewertet wurde als die Gruppe D.

Wenn wir die Differenz-Variablen beider Gruppen zum Vergleich heranziehen, erweisen sich deutliche Unterschiede. Die Tabelle der paarweisen Vergleiche der Differenz-Variablen im Anhang 9 zeigt für die Gruppen C und D ein angepasstes Signifikanzniveau von $p = 0,000$, das auf einen hochsignifikanten Unterschied zwischen diesen Variablen hinweist. Da die Gruppe C eine signifikant kleinere Differenz zu ihren parallelen korrekten Sätzen aufzeigt und die Gruppe D allgemein schlechter bewertet wurde, können wir an dieser Stelle die Hypothese 1.3 definitiv als falsifiziert betrachten.

4.2. Darstellung der Ergebnisse und Überprüfung der Hypothesen zu den syntaktischen und morphosyntaktischen Fehlerarten

Die nächsten Kapitel stellen zunächst die gewonnenen deskriptiven Daten über die Bewertung der syntaktischen und morphosyntaktischen Fehlerarten dar. Um einen Überblick über die Bewertungsunterschiede zwischen den grammatikalischen Fehlerarten zu gewinnen, vergleichen wir sie anhand einer Varianzanalyse. Bestätigt diese Analyse einen signifikanten Unterschied zwischen mindestens zwei Fehlerarten, führen wir im darauf folgenden Kapitel eine Folgeanalyse durch und versuchen dadurch herauszufinden, wie die einzelnen Fehlergruppen sich voneinander unterscheiden. Diese Analyse sollte uns ermöglichen, die in den Hypothesen 2.1, 2.2 und 2.3 aufgestellten Annahmen zu überprüfen.

4.2.1. Deskriptive Werte für die Beurteilung der syntaktischen und morphosyntaktischen Fehler

4.2.1.1. Fehlergruppe G: Genus

Die eingeplanten Genusfehler präsentierten wir durch die kursiv geschriebenen Wörter in folgenden Sätzen.

1.G. Ich habe leider *keinen* Zeit.

7.G. *Mein* Freundin ist naiv, aber nicht dumm.

13.G. Wie lange kannst du *deine* Atem anhalten?

19.G. Die Gastgeber servierten *eine* üppige Mahl.

25.G. Deutsch ist für mich *ein* zweites Sprache.

31.G. Wir treffen uns an *der* Hauptbahnhof.

In diesen Beispielen wurde der Genusfehler in verschiedenen Ausprägungen, nämlich als Negativartikel, Possessivartikel, bestimmter und unbestimmter Artikel, in verschiedenen Kasusformen also in Nominativ- bzw. Akkusativkontexten sowie nach der Präposition präsentiert. Die Tabelle (40) veranschaulicht, wie die einzelnen Sätze dieser Gruppe insgesamt bewertet wurden.

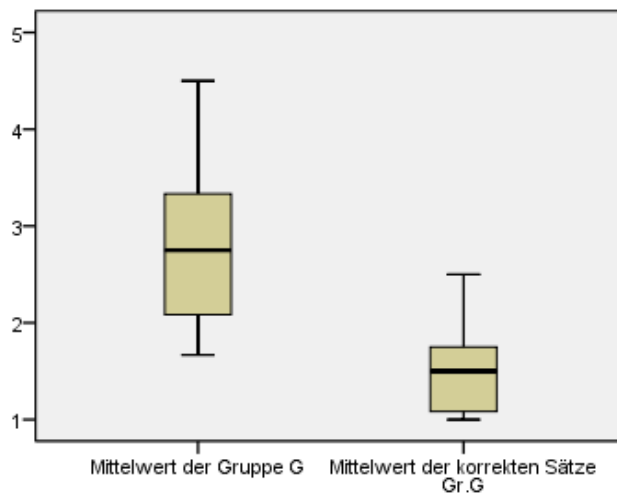
(40) Statistische Kennwerte für die Fehlergruppe G

		Satz1.G	Satz7.G	Satz13.G	Satz19.G	Satz25.G	Satz31.G
N	gültig	32	32	29	32	30	31
	fehlend	0	0	3	0	2	1
Mittelwert		2,594	2,063	2,483	3,375	3,533	2,806
Median		2,500	2,000	2,000	3,000	3,000	3,000
Modalwert		2,0	2,0	2,0	3,0	3,0	2,0
Standardabweichung		1,1601	,9136	,9495	1,0999	1,1059	1,0139

Wie die Tabelle zeigt, wurden die Sätze 19.G und 25.G mit einem höheren Mittelwert bzw. schlechter beurteilt als die anderen Sätze. Ein möglicher Grund wäre, dass in diesen Sätzen zwei Endungen, nämlich die Endung nach dem unbestimmten Artikel und die nach dem Adjektiv, fehlerhaft gesprochen wurden. Bei vier anderen Sätzen wurde nur eine Endung falsch gesprochen und sie wurden deshalb als weniger störend bewertet. Die absolute Übereinstimmung zwischen den drei Maßzahlen der zentralen Tendenz war bei keinem Satz zu beobachten. Dies reflektierte sich in den Ergebnissen des K-S-Tests, die auf eine Abweichung von der Normalverteilung der Bewertungen aller Sätze hinwiesen (s. Anhang 5.A.1).

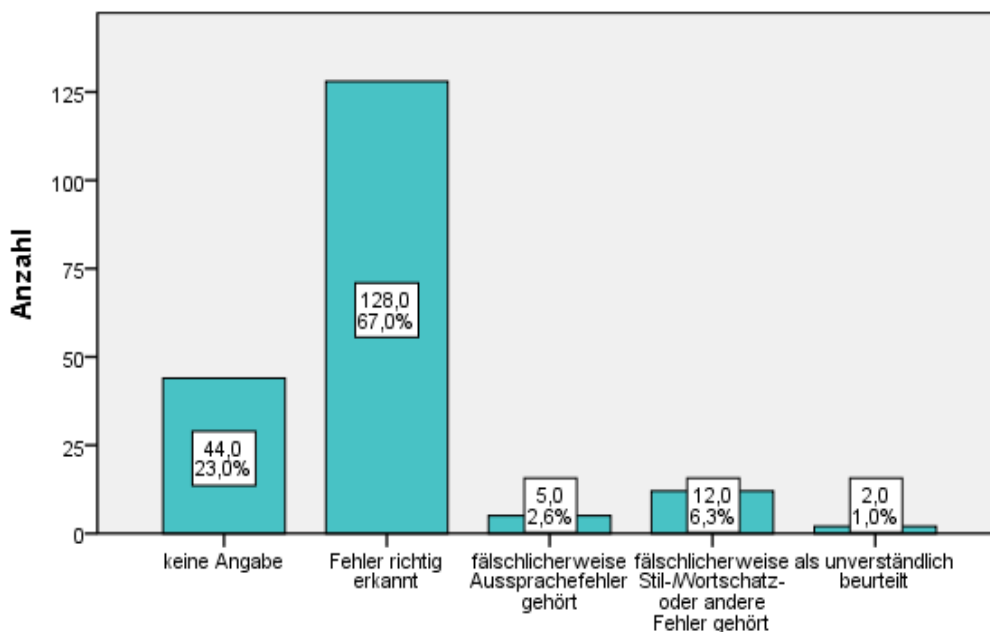
Die Gruppe G erzielte einen Gesamtmittelwert von ($M = 2,78$, $SD = 0,80$), der im Vergleich zu der parallelen Gruppe der korrekten Sätze ($M = 1,52$, $SD = 0,46$) deutlich höher lag. Für die Überprüfung der Bedeutsamkeit dieses Unterschieds wurde ein t-Test durchgeführt, dessen Ergebnis darauf hinwies, dass die Genusfehler im Vergleich zu den parallelen fehlerfreien Sätzen signifikant schlechter beurteilt wurden ($t(31) = 12,726$, $p = 0,000 < 0,01$, $r = 0,91$). Die unterschiedliche Lage der Mediane der beiden Gruppen, die im folgenden Boxplot-Diagramm veranschaulicht wird, bestätigt diese Annahme.

(41) Vergleichende Darstellung des Medians und der Streuung für die Gruppe G und die Gruppe mit parallelen korrekten Sätzen



Wenn wir die Ergebnisse der Analyse der Anmerkungen zu den erkannten Fehlern, die in folgender Tabelle präsentiert wird, betrachten, können wir sehen, dass 67% der Probanden, davon 68% mit dem Sicherheitsgrad 1 (s. Anhang 7), den Genusfehler im Satz richtig erkannt haben. Wie die folgende Übersicht zeigt, hat nur ca. 10% fälschlicherweise andere Fehlerarten angegeben, was darauf hinweist, dass der Genusfehler für deutsche Muttersprachler leicht zu erkennen und auffällig zu sein scheint.

(42) Fehlerangaben der Probanden für die Gruppe G



Um zu sehen, wie die richtige Erkennung des Fehlers in den einzelnen Sätzen aussah, betrachten wir die untere Tabelle.

(43) Absolute und relative Häufigkeit der korrekten Erkennung des Fehlers für die Sätze der Gruppe G

	Satz1.G	Satz 7.G	Satz 13.G	Satz 19.G	Satz 25.G	Satz31.G
Fehler richtig erkannt	24 75%	15 46,9%	9 28,1%	27 84,4%	24 75%	29 90,6%

Es fällt auf, dass der Genusfehler nach der Präposition im Satz 31.G von ca. 90% der Probanden, d. h. mehr als bei anderen Sätzen, erkannt wurde, während dieser Satz bei der Skalenbewertung (s. Tabelle 40) als weniger störend bzw. mit einem niedrigeren Mittelwert als die Sätze 19.G und 25.G bewertet worden war. Die Tatsache, dass manche Fehler sehr leicht erkannt, trotzdem als nicht störend bewertet werden, bildet das Thema der Hypothese 5, auf das wir an späterer Stelle detailliert eingehen werden.

4.2.1.2. Fehlergruppe H: Kasus

Bei der Konzeption der Beispiele für die fehlerhafte Kasusflexion beschränkten wir uns in dieser Arbeit auf die falsche Realisierung der Akkusativ- und Dativ-Ergänzungen.

- 2. H. Das Justizlexikon steht *ins* Regal.
- 8. H. Der Wind blies mir *im* Gesicht.
- 14. H. Sie ist die gute Seele in *unser* Team.
- 20. H. Kein Problem. Das nehme ich *dich* nicht übel.
- 26. H. Wir sind auf *die* Straße.
- 32. H. Die Kinder sind auf *den* Spielplatz.

Die kursiv geschriebenen Beispiele bestanden hauptsächlich aus einer falschen Realisierung des Kasus-Signals in Artikelwörtern nach einer Wechselpräposition. Der einzige Satz, der die fehlerhafte Kasus-Ergänzung des Verbs präsentierte, war der Satz 20.H, in dem das Akkusativpronomen *dich* anstelle des Dativpronomens *dir* gesprochen wurde. Die folgende Tabelle zeigt die deskriptiven Kennzahlen dieser Sätze.

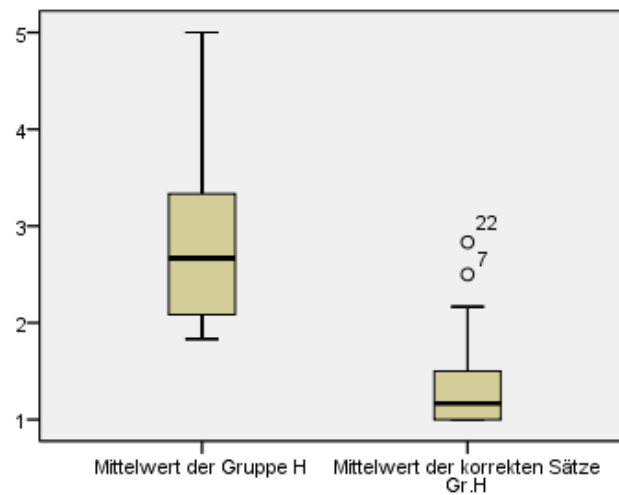
(44) Statistische Kennwerte für die Fehlergruppe H

	Satz2.H	Satz8.H	Satz14.H	Satz20.H	Satz26.H	Satz32.H
N gültig	32	32	32	32	32	29
fehlend	0	0	0	0	0	3
Mittelwert	3,344	2,906	2,656	3,031	3,156	1,759
Median	3,500	3,000	2,000	3,000	3,000	1,000
Modalwert	4,0	2,0	2,0	3,0	3,0	1,0
Standardabweichung	1,1807	1,0883	,8273	1,0313	1,0809	,9876

Wie wir sehen, wurde der Satz 32.H mit einem deutlich niedrigeren Mittelwert bewertet als alle anderen Sätze. Ob dies mit der ähnlichen Artikulationsart der Nasale [n] und [m] am Silbenende zusammenhängt oder auf reinen Zufall bzw. auf eine unauffällig gesprochene Sprechweise unseres Sprechers zurückzuführen ist, lässt sich an dieser Stelle schwer erklären. Bei den anderen Sätzen können wir davon ausgehen, dass sie relativ einheitlich mit den Mittelwerten zwischen 2,6 und 3,3 bewertet wurden. Auch in dieser Gruppe kann man bei fast keinem Satz eine Übereinstimmung der Kennzahlen der zentralen Tendenz feststellen und sie wiesen alle eine Abweichung von der Normalverteilung auf (s. Anhang 5.A.1).

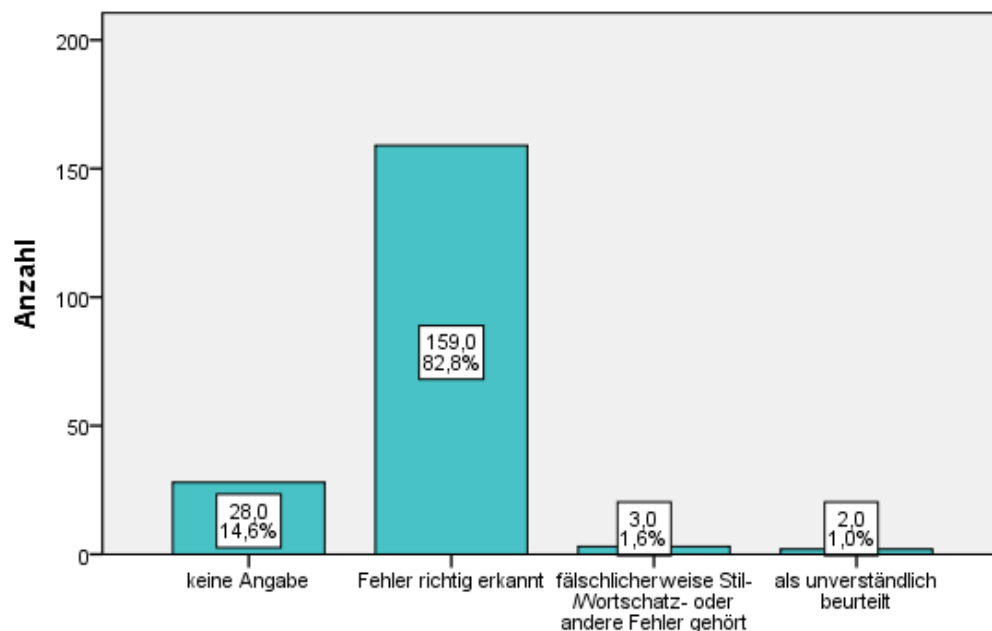
Die Gruppe der Kasusfehler wurde mit einem Gesamtmittelwert von ($M = 2,83$, $SD = 0,87$) deutlich schlechter bewertet als die parallele Gruppe mit korrekten Sätzen ($M = 1,37$, $SD = 0,47$). Da die Gruppe mit korrekten Sätzen keine Normalverteilung ihrer Daten aufwies, führten wir einen Wilcoxon-Test durch, dessen Ergebnis ($z = -4,941$, $p = 0,000 < 0,01$, $r = -0,87$) einen hochsignifikanten Unterschied zwischen diesen zwei Gruppen bestätigte. Das folgende Boxplot-Diagramm zeigt uns auch, wie auffällig schlechter die Gruppe der Kasusfehler im Vergleich zu der Gruppe der korrekten Sätze beurteilt wurde.

(45) Vergleichende Darstellung des Medians und der Streuung für die Gruppe H und die Gruppe mit parallelen korrekten Sätzen



Betrachten wir die folgende Übersicht zur Analyse der Anmerkungen der Probanden zu den erkannten Fehlern, sehen wir, dass hier ein großer Anteil der Probanden, nämlich 82,8%, davon ca. 72% mit der Sicherheitsangabe 1 (s. Anhang 7), den Kasusfehler im Satz erkannt hat. Hier haben nur ca. 3% der Probanden andere Fehlerarten im Satz entdeckt, was ein Zeichen dafür sein kann, dass der Kasusfehler für die meisten Deutschen deutlich hörbar und zu erkennen ist.

(46) Fehlerangaben der Probanden für die Gruppe H



Wenn wir die Prozentzahl der richtigen Erkennung in den einzelnen Sätzen in der unteren Tabelle betrachten, finden wir auch hier einen eindeutigen Unterschied zwischen dem Satz 32.H und den anderen Sätzen.

(47) Absolute und relative Häufigkeit der korrekten Erkennung des Fehlers für die Sätze der Gruppe H

	Satz 2.H	Satz 8.H	Satz 14.H	Satz 20.H	Satz 26.H	Satz 32.H
Fehler richtig erkannt	31 96,9%	29 90,6%	29 90,6%	30 93,8%	27 84,4%	13 40,6%

4.2.1.3. Fehlergruppe I: Präposition

Die Anwendung der falschen Präposition wurde anhand der kursiv geschriebenen Präpositionen in folgenden Sätzen dargestellt.

3.I. Er denkt nicht *über* seine Zukunft.

9.I. Er beteiligt sich immer aktiv *zum* Unterricht.

15.I. Mein Zahn ist sehr empfindlich und tut *am* Kauen weh.

21.I. Sie sind sehr gütig *mit* mir.

27.I. Ich habe Angst *gegen* Schlangen.

33.I. Meine Freundin ist immer noch *auf* dem Krankenhaus

Bei der Auswahl der Beispiele haben wir darauf geachtet, fehlerhafte Präpositionen auszuwählen, die einerseits im Satz inhaltlich Sinn machen, andererseits von DaF-Lernern nach meiner eigenen Erfahrung als Lehrkraft sehr oft verwendet werden. Wie die Beispiele zeigen, wurden die falschen Präpositionen in verschiedenen Ausprägungen, nämlich als Ergänzung des Verbs (Sätze 3.I und 9.I), des Adjektivs (Satz 21.I), des Nomens (Satz 27.I), als temporale (Satz 15.I) sowie als lokale Präposition (Satz 33.I) präsentiert. Die folgende Tabelle stellt die deskriptiven Kennzahlen der Sätze dieser Gruppe dar.

(48) Statistische Kennwerte für die Fehlergruppe I

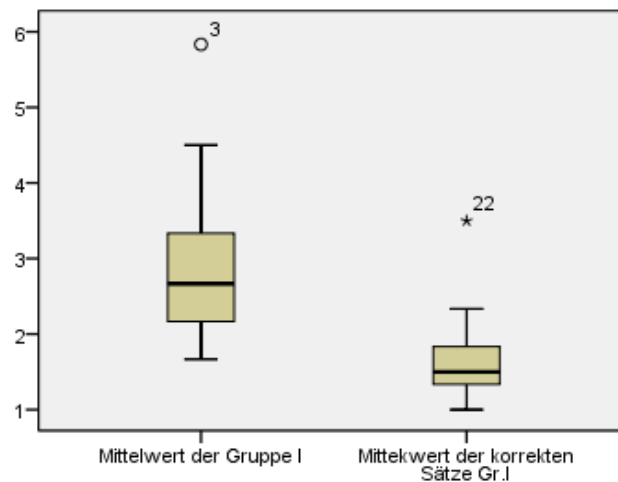
		Satz3.I	Satz9.I	Satz15.I	Satz21.I	Satz27.I	Satz33.I
N	gültig	32	31	29	28	32	32
	fehlend	0	1	3	4	0	0
Mittelwert		3,031	2,871	3,345	1,786	3,188	2,891
Median		3,000	3,000	3,000	1,000	3,000	3,000
Modalwert		3,0	2,0	3,0	1,0	3,0	3,0
Standardabweichung		,9995	1,0565	1,1425	1,0313	1,0298	1,1621

Wie aus der Tabelle ersichtlich ist, weist der Satz 21.I einen eindeutig niedrigeren Mittelwert im Vergleich zu den anderen Sätzen auf. Dies könnte damit begründet werden, dass das Adjektiv *gütig* in Verbindung mit mehreren Präpositionen, nämlich mit *gegen*, *gegenüber* und *zu*, angewendet werden kann. Diese Auswahl führt die Probanden vermutlich dazu, die Verwendung anderer Präpositionen für möglich zu halten, weshalb sie die Präposition *mit* als weniger fehlerhaft bewertet haben. Eine solche Interpretation bildet allerdings eine eigene Hypothese, die durch eine empirische Analyse überprüft werden müsste.

Obwohl bei fast allen Sätzen die Kennwerte von Median und Modalwert identisch sind, ist zwischen diesen Kennzahlen und dem Mittelwert keine Übereinstimmung aufzuweisen. Diese Tatsache beruht auf der Existenz der Ausreißer, die bei allen Sätzen zu einer Abweichung der Bewertungen von der Normalverteilung geführt haben (s. Anhang 5.A.1).

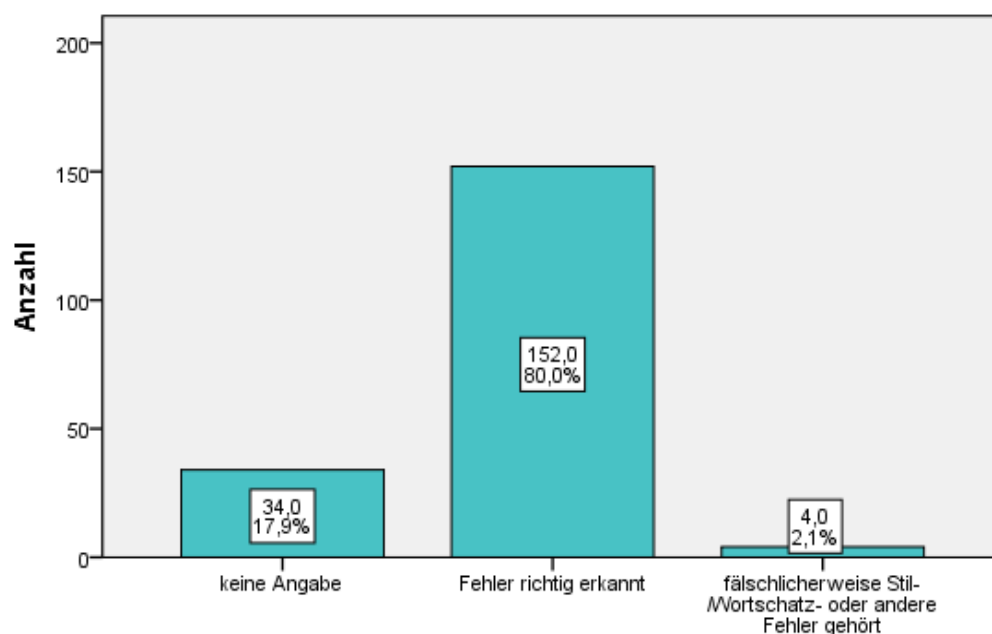
Für die Gruppe mit Präpositionsfehlern wurde insgesamt ein Mittelwert von ($M = 2,86$, $SD = 0,90$) berechnet, der im Vergleich zu der Gruppe mit parallelen fehlerfreien Sätzen ($M = 1,60$, $SD = 0,51$) deutlich höher lag. Das Ergebnis des Wilcoxon-Tests zeigte auf, dass die Gruppe I signifikant schlechter beurteilt wurde als die Gruppe der korrekten Sätze ($z = -4,895$, $p = 0,000 < 0,01$, $r = -0,86$). Die unterschiedliche Lage des Medians dieser zwei Gruppen ist am folgenden Boxplot-Diagramm leicht zu erkennen.

(49) Vergleichende Darstellung des Medians und der Streuung für die Gruppe I und die Gruppe mit parallelen korrekten Sätzen



In dieser Gruppe hat, wie in folgender Übersicht zu sehen ist, ein großer Anteil bzw. 80% der Probanden, davon 73% mit dem Sicherheitsgrad 1 (s. Anhang 7), die falsche Präposition im Satz als Fehler erkannt. Auch hier hat nur ein geringer Anteil der Probanden, nämlich ca. 2% andere Fehlerarten als Fehler angemerkt, was darauf hinweist, dass die falsche Präposition im Satz für deutsche Muttersprachler leicht zu erkennen sein muss.

(50) Fehlerangaben der Probanden für die Gruppe I



Um zu sehen, wie hoch die Anzahl der erkannten Fehler in den einzelnen Sätzen aussah, betrachten wir die Tabelle (51). Auch hier fällt auf, dass beim Satz 21.I weniger Probanden die falsche Präposition *mit* für das Adjektiv *gütig* als Fehler erkannt haben.

(51) Absolute und relative Häufigkeit der korrekten Erkennung des Fehlers für die Sätze der Gruppe I

	Satz 3.I	Satz 9.I	Satz 15.I	Satz 21.I	Satz 27.I	Satz33.I
Fehler richtig erkannt	28 87,5%	28 87,5%	29 90,6%	8 25%	31 96,9%	28 87,5%

4.2.1.4. Fehlergruppe J: Partizip II

Die in den folgenden Sätzen markierten Wörter präsentierten den morphosyntaktischen Fehler bei der Produktion des Partizips II.

4.J. Meine Augen waren vom Glanz der Sonne *geblenden*.

10.J. Einer der Hunde war sehr aggressiv und hat andere Hunde *verletzen*.

16.J. Es wird ewig dauern, bis er seine Schulden *bezahlen* hat.

22.J. Sie hat sich gar nicht *bemühen*.

28.J. Die Verbrennungen haben ihr unerträgliche Schmerzen *bereiten*.

34.J. Der Taxifahrer hat an der Bushaltestelle *gehalten*.

Bei der Auswahl der Partizipien haben wir darauf geachtet, dass nur das Suffix des Partizips bzw. die Endung falsch gesprochen wird und der Stammvokal sowie das Präfix unverändert bleiben. Wie wir sehen, bestanden unsere Beispiele hauptsächlich aus regelmäßigen Verben mit dem Suffix *-t*, das fälschlicherweise als *-en* gesprochen wurde. Das einzige unregelmäßige Verb wurde im Satz 34.J präsentiert, wobei auch bei diesem Verb keine Änderung des Stammvokals vorlag. Die folgende Übersicht zeigt uns, wie die einzelnen Sätze bewertet wurden.

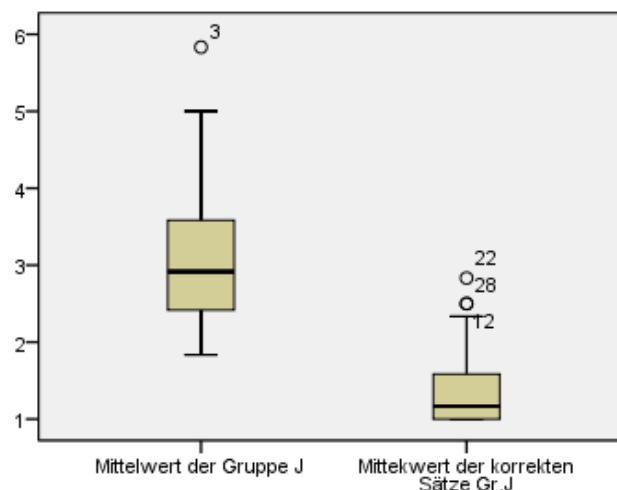
(52) Statistische Kennwerte für die Fehlergruppe J

		Satz4.J	Satz10.J	Satz16.J	Satz22.J	Satz28.J	Satz34.J
N	gültig	31	30	32	30	32	32
	fehlend	1	2	0	2	0	0
Mittelwert		2,935	3,400	2,813	3,367	2,813	3,500
Median		3,000	3,000	3,000	3,000	3,000	4,000
Modalwert		2,0	3,0	3,0	3,0	2,0	4,0
Standardabweichung		1,1236	1,0372	1,0298	,9994	,9311	1,2181

Wie wir sehen, unterscheidet sich die Anzahl der zentralen Tendenzen der Sätze in der Gruppe J jeweils minimal voneinander. In dieser Gruppe wurde kein Satz signifikant besser oder schlechter beurteilt als die anderen Sätze. Die Übereinstimmung zwischen den Kennzahlen der mittleren Tendenz ist bei keinem Satz festzustellen und die Sätze wiesen deshalb keine normalverteilten Bewertungen auf (s. Anhang 5.A.1).

Die Sätze mit Partizip-II-Fehlern wurden insgesamt mit dem Mittelwert von ($M = 3,12$, $SD = 0,93$) deutlich schlechter bewertet als die parallelen fehlerfreien Sätze ($M = 1,42$, $SD = 0,52$). Der Wilcoxon-Test bestätigte diesen Unterschied als hochsignifikant ($z = -4,941$ $p = 0,000 < 0,01$, $r = -0,87$). Das folgende Boxplot-Diagramm veranschaulicht diesen Unterschied durch die Darstellung der Mediane der beiden Gruppen.

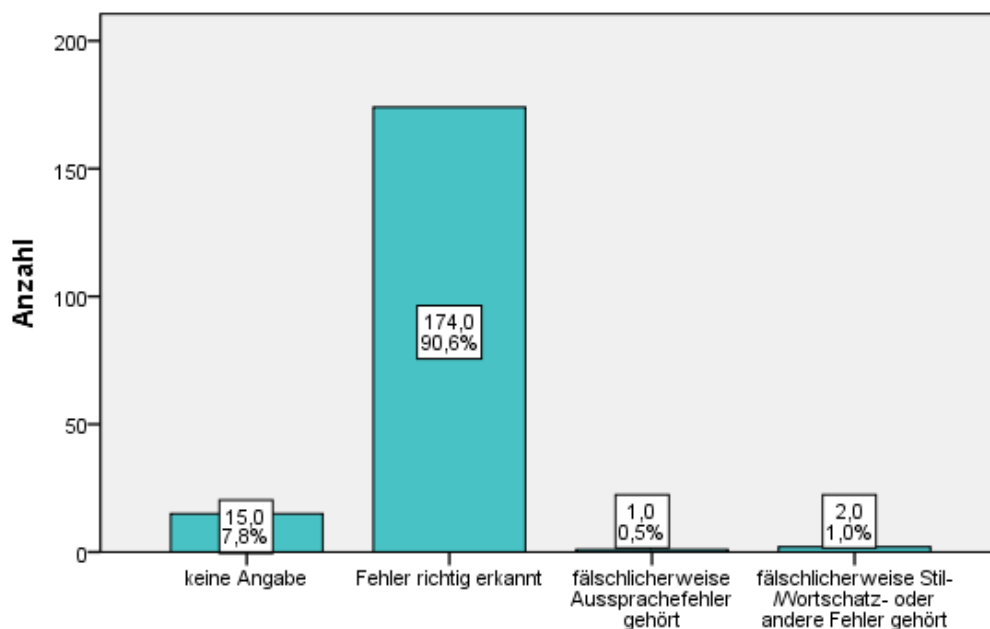
(53) Vergleichende Darstellung des Medians und der Streuung für die Gruppe J und die Gruppe mit parallelen korrekten Sätzen



Die Fehlerangaben in der Abbildung (54) zeigen, dass die falsche Partizip-II-Form von 90,6% der Probanden, davon ca. 74% mit dem Sicherheitsgrad 1 (s. Anhang 7), d. h. mehr als

bei allen bis jetzt untersuchten Fehlerarten, als Fehler erkannt wurde. Nur ca. 2% der Probanden haben andere Fehlerarten angegeben, was zeigt, dass diese Fehlerart für deutsche Muttersprachler auffällig hörbar ist.

(54) Fehlerangaben der Probanden für die Gruppe J



Wenn wir die folgende Übersicht betrachten, können wir feststellen, dass die Fehler in den einzelnen Sätzen dieser Gruppe sehr einheitlich und mit einer hohen Prozentzahl erkannt wurden.

(55) Absolute und relative Häufigkeit der korrekten Erkennung des Fehlers für die Sätze der Gruppe J

	Satz 4.J	Satz 10.J	Satz 16.J	Satz 22.J	Satz 28.J	Satz34.J
Fehler richtig erkannt	29	28	29	29	29	30
	90,6%	87,5%	90,6%	90,6%	90,6%	93,8%

4.2.1.5. Fehlergruppe K: Verbposition im Hauptsatz

Die falsche Stellung des Verbs im Hauptsatz wurde anhand folgender Beispiele präsentiert.

5 .K. Neulich Benzin *ist* viel teurer geworden.

11.K. Leider die Rechnung *wies* einige Fehler auf.

17.K. Am Wochenende das Wetter *sollte* schön werden.

23.K. Jeden Tag die Künstler *treffen sich*.

29.K. Letztes Jahr ich *habe* mein Studium beendet.

35.K. Als ausländischer Student man *muss* in Deutschland zuerst einen Sprachkurs besuchen.

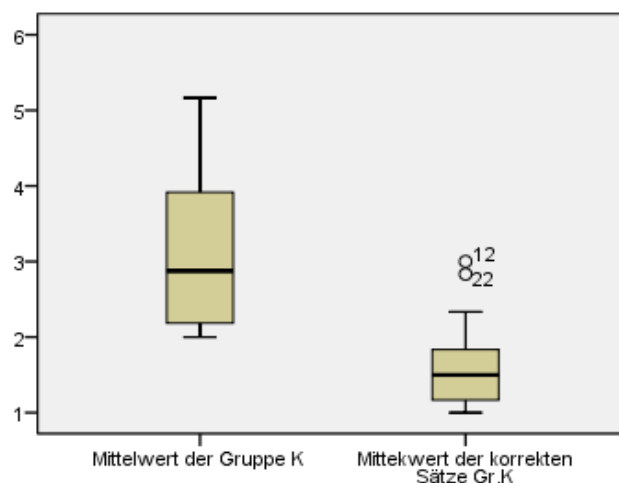
Diese Sätze beruhen auf der Verletzung der Umstellungsregel, die besagt: Wenn ein anderes Satzglied in der Position I steht, dann folgen das konjugierte Verb in der zweiten und das Subjekt in der dritten Position. Vielen DaF-Lernern kommt die Positionierung des Subjekts nach dem Verb bzw. die Umstellungsregel seltsam vor, weshalb nach meiner eigenen Erfahrung als DaF-Lehrkraft viele diese Regel verletzen. In folgender Tabelle können wir sehen, wie die einzelnen Sätze dieser Fehlergruppe bewertet wurden.

(56) Statistische Kennwerte für die Fehlergruppe K

		Satz5.K	Satz11.K	Satz17.K	Satz23.K	Satz29.K	Satz35.K
N	gültig	31	30	32	31	32	31
	fehlend	1	2	0	1	0	1
Mittelwert		3,484	3,233	3,031	3,290	3,000	2,839
Median		3,000	3,000	3,000	3,000	3,000	3,000
Modalwert		2,0	2,0	2,0	2,0	2,0	2,0
Standardabweichung		1,2348	1,2507	1,1496	1,0706	1,0473	,9344

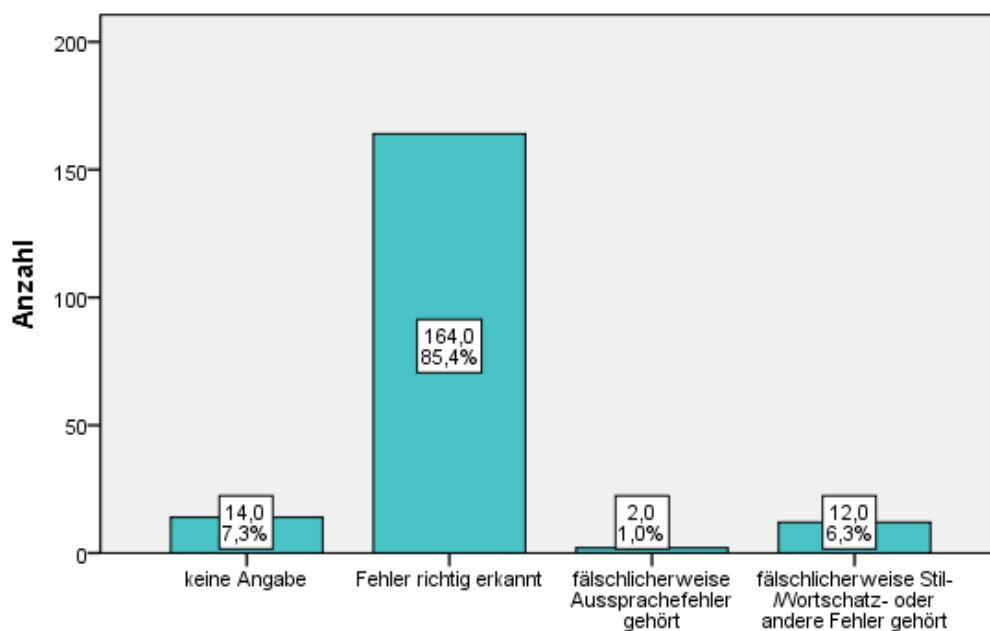
Wie die Tabelle zeigt, wurden die Sätze der Gruppe K relativ einheitlich mit Mittelwerten zwischen 2,8 und 3,4 bewertet. Auch hier ist bei keinem Satz eine Übereinstimmung zwischen den Maßzahlen der zentralen Tendenz zu beobachten und die Sätze wurden alle laut K-S-Test als nicht normalverteilt eingestuft (s. Anhang 5.A.1). Die Gruppe K wies einen Gesamtmittelwert von ($M = 3,13$, $SD = 0,99$) auf, der im Vergleich zu der Gruppe mit parallelen korrekten Sätzen ($M = 1,54$, $SD = 0,52$) eindeutig höher lag. Dieser Unterschied wurde auch durch einen Wilcoxon-Test als hochsignifikant ($z = -4,941$, $p = 0,000 < 0,01$, $r = -0,87$) erwiesen. Die unterschiedliche Lage der Mediane, die in folgender Übersicht zu sehen ist, bestätigt diesen Unterschied.

(57) Vergleichende Darstellung des Medians und der Streuung für die Gruppe K und die Gruppe mit parallelen korrekten Sätzen



In der Gruppe K hat ein großer Anteil der Probanden, nämlich 85,4%, davon 71% mit der Sicherheitsangabe 1 (s. Anhang 7), die falsche Position des Verbs im Hauptsatz als Fehler angemerkt. Auch hier hat nur ein geringer Anteil andere Fehlerarten angegeben, was darauf zurückzuführen ist, dass diese Fehlerart für deutsche Hörer sehr auffällig klingt.

(58) Fehlerangaben der Probanden für die Gruppe K



Die Tabelle (59) zeigt, wie die falsche Verbposition im Hauptsatz in den einzelnen Sätzen als Fehler erkannt wurde. Auch hier, wie bei den Bewertungsergebnissen in der Tabelle (56),

weist kein Satz einen auffälligen Unterschied der korrekten Erkennung des Fehlers im Vergleich zu den anderen Sätzen auf.

(59) Absolute und relative Häufigkeit der korrekten Erkennung des Fehlers für die Sätze der Gruppe K

	Satz 5.K	Satz 11.K	Satz 17.K	Satz 23.K	Satz 29.K	Satz35.K
Fehler richtig erkannt	24 75%	30 93,8%	26 81,3%	26 81,3%	29 90,6%	29 90,6%

4.2.1.6. Fehlergruppe L: Verbposition im Nebensatz

Die fehlerhafte Verbposition im Nebensatz wurde anhand folgender Sätze, in denen das finite Verb anstatt am Satzende nach dem Subjekt stand, dargestellt.

6. L. Alle wissen, dass die Touristen *haben* sich mit Malaria infiziert.

12. L. Wenn du *willst* dich mehr informieren, lies dieses Buch.

18. L. Die Kinder spielen mit dem Ball, obwohl sie *dürfen* es gar nicht.

24. L. Auch wenn es hier *ist* sehr gemütlich, müssen wir bald wieder aufbrechen.

30. L. Obwohl er *ist* nicht so groß, kann er drei Meter hoch springen.

36. L. Ich wusste nicht, dass das Studienkolleg *ist* hier in der Nähe.

Bei der Auswahl dieser Beispiele beschränkten wir uns auf Nebensätze, die mit den Konjunktionen *dass*, *obwohl* und *wenn* eingeleitet wurden. Auf kausale Nebensätze mit *weil* haben wir bewusst verzichtet, da das Verb in Nebensätzen mit dieser Konjunktion in der Umgangssprache häufig wie im Hauptsatz, d. h. nach dem Subjekt und nicht in der Endposition, gesprochen wird. Die Tabelle unten stellt eine Übersicht der quantitativen Informationen über die Bewertung der einzelnen Sätze dieser Gruppe dar.

(60) Statistische Kennwerte für die Fehlergruppe L

	Satz6.L	Satz12.L	Satz18.L	Satz24.L	Satz30.L	Satz36.L
N gültig	32	32	31	31	31	32
fehlend	0	0	1	1	1	0
Mittelwert	3,531	3,438	3,226	3,000	3,323	3,281
Median	3,000	3,000	3,000	3,000	3,000	3,000
Modalwert	3,0	3,0	3,0	2,0	2,0	3,0
Standardabweichung	1,1067	1,1053	1,1750	,9309	1,0766	1,0846

In dieser Gruppe gab es keinen Satz, dessen Mittelwert augenfällig höher oder niedriger war als die anderen, und wie wir sehen, wurden die Sätze sehr einheitlich mit Mittelwerten zwischen 3 und 3,5 bewertet. Eine Übereinstimmung zwischen den Maßzahlen der zentralen Tendenz war bei keinem Satz zu beobachten und die Kurve der Bewertungen bei allen Sätzen zeigte eine Abweichung von der Normalverteilung auf.

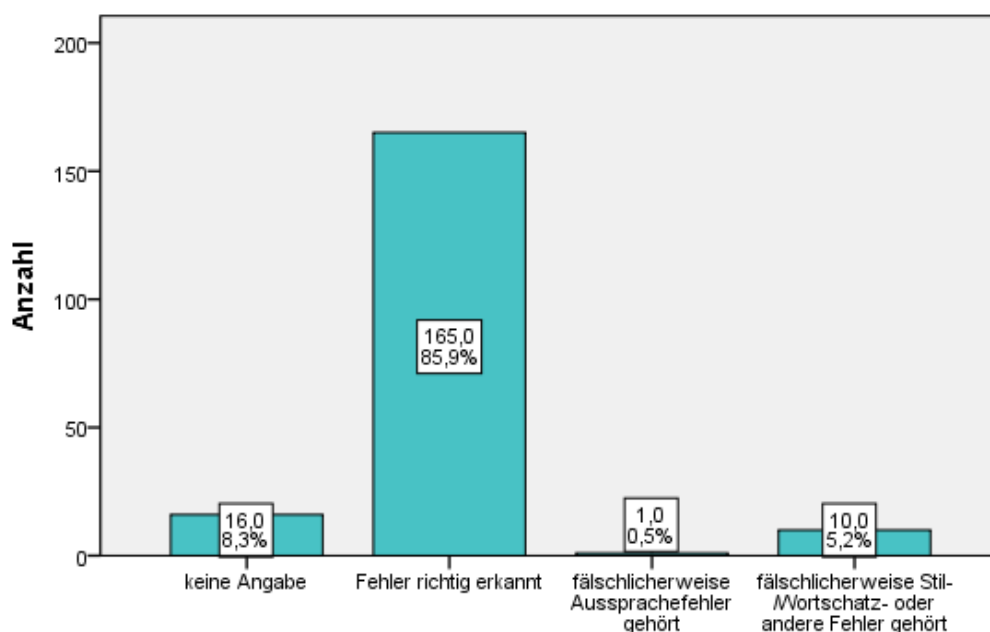
Die Gruppe L wurde insgesamt mit einem Mittelwert von ($M = 3,29$, $SD = 0,93$) bewertet, der im Vergleich zu der Gruppe mit parallelen korrekten Sätzen ($M = 1,47$, $SD = 0,46$) deutlich höher war. Durch einen Wilcoxon-Test wurde bestätigt, dass die Gruppe mit Fehlern in der Positionierung des Verbs im Nebensatz signifikant schlechter bewertet wurde als die Gruppe mit parallelen korrekten Sätzen ($z = -4,944$, $p = 0,000 < 0,01$, $r = -0,87$). Die unterschiedliche Position der Mediane dieser zwei Gruppen, die anhand folgenden Boxplot-Diagramms veranschaulicht wird, bestätigt diese Annahme.

(61) Vergleichende Darstellung des Medians und der Streuung für die Gruppe L und die Gruppe mit parallelen korrekten Sätzen



Die falsche Verbposition im Nebensatz wurde, wie das folgende Diagramm zeigt, von ca. 86% der Probanden, davon ca. 68% mit der Sicherheitsangabe 1 (s. Anhang 7), als Fehler gekennzeichnet.

(62) Fehlerangaben der Probanden für die Gruppe L



Betrachten wir die Anzahl der korrekten Erkennung bei den einzelnen Sätzen, können wir feststellen, dass die syntaktischen Fehler in der Positionierung des Verbs im Nebensatz mit verschiedenen Konjunktionen sehr einheitlich und von der Mehrheit der Probanden erkannt wurden.

(63) Absolute und relative Häufigkeit der korrekten Erkennung des Fehlers für die Sätze der Gruppe L

	Satz 6.L	Satz 12.L	Satz 18.L	Satz 24.L	Satz 30.L	Satz36.L
Fehler richtig erkannt	28	25	28	27	29	28
	87,5%	78,1%	87,5%	84,4%	90,6%	87,5%

4.2.2. Statistische Auswertung der gewonnenen Daten über die Bewertung der syntaktischen und morphosyntaktischen Fehlerarten

4.2.2.1. Varianzanalyse und die Rangordnung der Grammatik-Fehlergruppen

Wie im Kapitel 4.1.2.1. begründet, führen wir auch hier für die grammatikalischen Fehlerarten aufgrund der Verborgenheit unserer unabhängigen Variablen eine nichtparametrische Varianzanalyse für mehr als zwei abhängige Stichproben, nämlich den Friedman-Test durch.

Laut den Ergebnissen dieser Analyse ($\chi^2(5) = 43,744$, $p = 0,000 < 0.01$) unterscheiden sich mindestens zwei Fehlergruppen signifikant voneinander. Die untere Tabelle präsentiert, wie der mittlere Rang bei den einzelnen Gruppen aussah.

(64) Friedman-Test: Ranginformation zu den Grammatik-Fehlergruppen

	Mittlerer Rang
Mittelwert der Gruppe G	2,45
Mittelwert der Gruppe H	2,72
Mittelwert der Gruppe I	2,81
Mittelwert der Gruppe J	4,27
Mittelwert der Gruppe K	4,08
Mittelwert der Gruppe L	4,67

Wie bei den phonologischen Fehlergruppen verwenden wir die oben dargestellten Ergebnisse als eine Basis-Information, um unsere Fehlerarten hinsichtlich ihrer Wirkung auf die deutschen Muttersprachler zu klassifizieren. Ordnen wir die Rangplätze der Größe nach und teilen ihnen die Plätze 1-6 zu, ergibt sich die folgende Rangliste.

(65) Rangordnung der Grammatikfehler hinsichtlich ihrer Wirkung auf die deutschen Muttersprachler

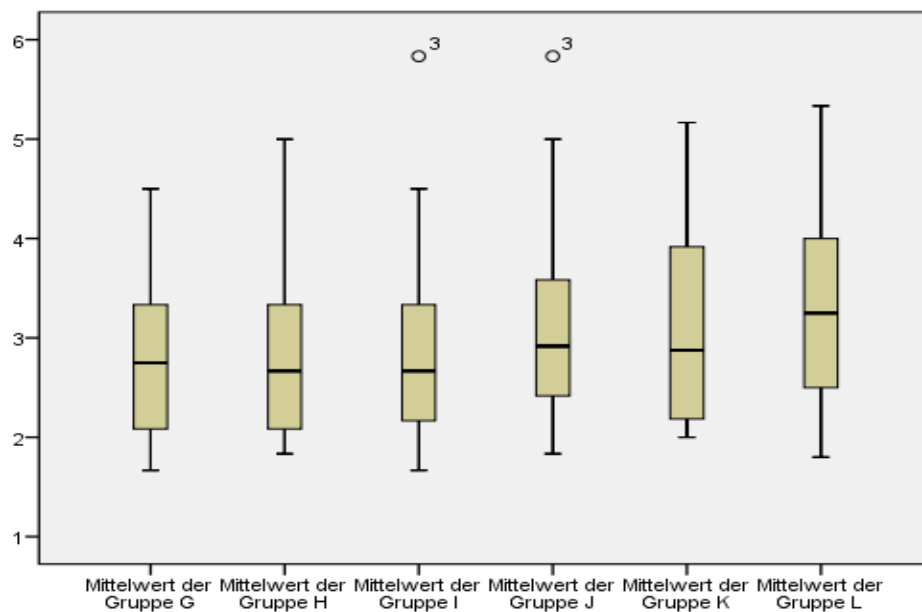
Platz	Gruppe	Mittlerer Rang
1	Gruppe L (Verbposition im Nebensatz)	4,67
2	Gruppe J (Partizip II)	4,27
3	Gruppe K (Verbposition im Hauptsatz)	4,08
4	Gruppe I (Präposition)	2,81
5	Gruppe H (Kasus)	2,72
6	Gruppe G (Genus)	2,45

Wie aus der Tabelle ersichtlich ist, wurden die ersten drei Fehlerarten von L1-Sprechern eindeutig weniger toleriert als die drei weiteren Gruppen. Der Unterschied zwischen den mittleren Rängen der drei Gruppen L, J und K sowie zwischen den drei anderen Gruppen, nämlich I, H und G, ist allerdings sehr gering. Anders als bei den phonologischen Fehlerarten bewegen sich hier die durchschnittlichen Ränge nur um zwei Zahlen, nämlich um 4 und 2. Aufgrund dieses Ergebnisses können wir davon ausgehen, dass die Grammatik-Fehlerarten sich hinsichtlich ihrer störenden Wirkung auf deutsche Muttersprachler in zwei Gruppen

aufteilen. Ob diese Aufteilung auf signifikanten Unterschieden basiert, zeigt sich anhand der Folgeanalyse, die wir im nächsten Kapitel durchführen werden.

An dieser Stelle betrachten wir die folgenden Boxplot-Diagramme, um zu sehen, ob wir die gleiche Rangreihe bei den Medianen feststellen können.

(66) Vergleichende Darstellung des Medians und der Streuung für die grammatischen Fehlerarten



Die Boxen, die 50% der Werte beinhalten, zeigen auch hier für die Gruppen J, K und L eine relativ höhere Lage im Vergleich zu den anderen Gruppen. Wenn wir die Position der Mediane betrachten, weisen die Gruppen J, K und L zwar höhere Mediane auf, aber der Unterschied ist hier nicht so auffällig wie bei der Ranginformation des Friedman-Tests. Eine Übersicht über die berechneten Mediane und Mittelwerte dieser sechs Gruppen kann im Anhang 8 eingesehen werden.

Wie bei den phonologischen Fehlergruppen vergleichen wir nun die Differenzen zwischen den Mittelwerten der grammatikalischen Fehlergruppen und der parallelen korrekten Sätze. Hierfür subtrahieren wir den Mittelwert der fehlerhaften Gruppe von dem Mittelwert der parallelen korrekten Sätze. Wenn wir die sechs neu entstandenen Differenz-Variablen durch einen Friedman-Test untersuchen, zeigt uns das Ergebnis ($\chi^2(5) = 32,274, p = 0,000 < 0.01$), dass auch hier mindestens zwei Gruppen signifikant unterschiedlich sind. Die Ranginformation dieses Tests zeigt eine ähnliche Reihenfolge wie die Tabelle (65).

(67) Friedman-Test: Ranginformation zu den Differenzen der grammatikalischen Fehlergruppen und parallelen korrekten Sätze

	Mittlerer Rang
Differenz G-KG	2,64
Differenz H-KH	3,30
Differenz I-KI	2,58
Differenz J-KJ	4,06
Differenz K-KK	3,75
Differenz L-KL	4,67

Die gleiche Rangreihe zeigt sich hier bei den Gruppen L, J und K. Die drei weiteren Gruppen weisen jedoch eine andere Reihenfolge auf als in der Tabelle (65), nämlich H gefolgt von G und I. Wenn wir die Prozentzahl der korrekten Erkennung und die dazugehörige Prozentzahl mit Sicherheitsgrad 1 in diesen Gruppen nebeneinander betrachten, stellt sich eine andere Rangliste heraus.

(68) Relative Häufigkeit der korrekten Erkennung der Grammatikfehler und die als ganz sicher angekreuzte Prozentzahl der korrekten Erkennung

Gruppen	G	H	I	J	K	L
Prozentzahl der korrekten Erkennung	67%	82,8%	80%	90,6%	85,4%	85,9%
Prozentzahl der mit Sicherheitsgrad 1 angekreuzten korrekten Erkennung	68%	72%	73%	74%	71%	68%

Wie wir sehen, wurden die Partizip-II-Fehler in der Gruppe J mehr als alle anderen Fehlerarten erkannt. Die Gruppe G wies auch hier die niedrigste Anzahl der korrekten Erkennung auf und die Fehler in den Gruppen L und K wurden besser erkannt als die in den Gruppen H und I.

Wenn wir die oben dargestellten Ergebnisse miteinander vergleichen, können wir die Schlussfolgerung ziehen, dass die drei Fehlerarten, nämlich die falsche Verbposition im Nebensatz (Gruppe L) und im Hauptsatz (Gruppe K) sowie Partizip-II-Fehler (Gruppe J), für deutsche Muttersprachler störender klingen als die Präpositions- (Gruppe I), Kasus- (Gruppe H) und Genusfehler (Gruppe G).

4.2.2.2. Paarvergleiche bzw. Post-hoc-Analyse der Grammatik-Fehlergruppen

Um die im letzten Kapitel gezogenen Schlussfolgerungen bzw. die Aufteilung der grammatischen Fehlergruppen hinsichtlich ihrer störenden Wirkung in zwei Gruppen auf Signifikanz zu überprüfen, führen wir in diesem Abschnitt eine Folgeanalyse durch und versuchen, mithilfe eines paarweisen Vergleichs herauszufinden, welche Fehlergruppen in unserer Rangliste sich tatsächlich bzw. signifikant voneinander unterscheiden. Die folgende Tabelle stellt die Ergebnisse dieser Paarvergleich-Analyse dar.

(69) Paarweiser Vergleich der Bewertung grammatischer Fehlerarten

Stichprobe1-Stichprobe2	Test-statistik	Std. Fehler	Standard Test-statistik	Sig.	Angep. Sig.
Mittelwert der Gruppe G-Mittelwert der Gruppe H	-,266	,468	-,568	,570	1,000
Mittelwert der Gruppe G-Mittelwert der Gruppe I	-,359	,468	-,768	,442	1,000
Mittelwert der Gruppe G-Mittelwert der Gruppe K	-1,625	,468	-3,474	,001	,008
Mittelwert der Gruppe G-Mittelwert der Gruppe J	-1,812	,468	-3,875	,000	,002
Mittelwert der Gruppe G-Mittelwert der Gruppe L	-2,219	,468	-4,744	,000	,000
Mittelwert der Gruppe H-Mittelwert der Gruppe I	-,094	,468	-,200	,841	1,000
Mittelwert der Gruppe H-Mittelwert der Gruppe K	-1,359	,468	-2,906	,004	,055
Mittelwert der Gruppe H-Mittelwert der Gruppe J	-1,547	,468	-3,307	,001	,014
Mittelwert der Gruppe H-Mittelwert der Gruppe L	-1,953	,468	-4,176	,000	,000
Mittelwert der Gruppe I-Mittelwert der Gruppe K	-1,266	,468	-2,706	,007	,102
Mittelwert der Gruppe I-Mittelwert der Gruppe J	-1,453	,468	-3,107	,002	,028
Mittelwert der Gruppe I-Mittelwert der Gruppe L	-1,859	,468	-3,976	,000	,001
Mittelwert der Gruppe K-Mittelwert der Gruppe J	,188	,468	,401	,688	1,000
Mittelwert der Gruppe K-Mittelwert der Gruppe L	-,594	,468	-1,269	,204	1,000
Mittelwert der Gruppe J-Mittelwert der Gruppe L	-,406	,468	-,869	,385	1,000

Wenn wir die als signifikant markierten Unterschiede in der letzten Reihe anschauen, können wir feststellen, dass diese nur zwischen den drei in der Tabelle (65) als weniger störend empfundenen Fehlerarten (Gruppen G, H und I) und den drei Gruppen mit gravierenderen

Fehlern (Gruppen J, K und L) zu finden sind. Der signifikante Unterschied zu allen drei Gruppen J, K und L zeigte sich allerdings nur bei der Gruppe G.

Werfen wir einen Blick auf die Ergebnisse der paarweisen Analyse der Differenz-Variablen im Anhang 9, sind zwar weniger signifikante Unterschiede zu beobachten, aber sie bestehen wiederum zwischen den Fehlergruppen (G, H und I) und den Gruppen (J und L).

Wenn wir diese Ergebnisse zusammen mit der Tatsache, dass es keine signifikanten Unterschiede zwischen den Gruppen (G, H und I) einerseits und den Gruppen (J, K und L) andererseits gibt, betrachten, bestätigt sich die Annahme, dass die in dieser Arbeit behandelten Grammatik-Fehlergruppen sich hinsichtlich ihrer Wirkung auf die Deutschen tatsächlich in zwei Gruppen aufteilen lassen. An dieser Stelle können wir die im letzten Kapitel gezogene Schlussfolgerung, dass die Fehler in der Positionierung des Verbs im Haupt- sowie Nebensatz und der Bildung des Partizips II für deutsche Muttersprachler weniger zu tolerieren sind als die falsche Realisierung des Genus, Kasus sowie der Präposition, auf signifikante Unterschiede zurückführen.

4.2.3. Überprüfung der Hypothesen zu den grammatischen Fehlern

Die durch die Paarvergleich-Analyse gewonnenen Erkenntnisse können wir nun zur Beantwortung der Fragen, die in den Hypothesen 2.1, 2.2 und 2.3 aufgeworfen wurden, einsetzen. Da es sich bei diesen drei Hypothesen hauptsächlich um den Vergleich von zwei Fehlergruppen handelt und wir die Fehler-Wahrscheinlichkeit, die bei mehrfachen Testwiederholungen zustande kommt, reduzieren möchten, beziehen wir uns für die Bestätigung oder Ablehnung der Hypothesen vor allem auf die Ergebnisse der Post-hoc-Analyse in der Tabelle (69). Wir vergleichen jedoch bei jeder Hypothese das Ergebnis dieser Analyse mit dem Ergebnis eines für den Vergleich von nur zwei Gruppen geeigneten Signifikanztests, um zu sehen, ob dort bemerkenswerte Unterschiede vorhanden sind.

4.2.3.1. Hypothese 2.1

In der Hypothese 2.1 wurde vermutet, dass die Genusfehler für die deutschen Probanden schwerwiegender wären als die falsche Realisierung des Kasus. Statistisch ausgedrückt, sollte der Gesamtmittelwert der Gruppe G (Genus) signifikant größer sein als die durchschnittliche Bewertung der Gruppe H (Kasus).

Betrachten wir die deskriptiven Ergebnisse dieser zwei Gruppen, ist der Mittelwert der Gruppe G ($M = 2,78$, $SD = 0,80$) sogar niedriger als der der Gruppe H ($M = 2,83$, $SD = 0,87$). Die dargestellten Medianpositionen im Diagramm (66) zeigen aber das Gegenteil. Wie das Boxplot-Diagramm zeigt, liegt der Median der Gruppe G ($Md = 2,75$) höher als der der Gruppe H ($Md = 2,66$). Da die beiden Gruppen laut den Ergebnissen des K-S-Tests (s. Anhang 5.B) als normalverteilt eingestuft wurden und unsere Variablen als intervallskaliert gelten, berücksichtigen wir hier für die Bestimmung der Lage der zentralen Tendenz eher das arithmetische Mittel. Der niedrigere Mittelwert der Gruppe G kann an dieser Stelle als Beweis dafür gelten, dass die Hypothese 2.1 nicht verifiziert werden kann.

Um zu sehen, ob der Unterschied zwischen diesen zwei Gruppen überhaupt signifikant ist, betrachten wir die Tabelle (69). Wie wir sehen, unterscheiden sich die Gruppen G und H mit einem angepassten P-Wert von 1,00 nicht signifikant voneinander. Hätten wir keine mehrfache Paarvergleich-Analyse vorgenommen, sondern einen t-Test durchgeführt, wären wir auch zu einem unsignifikanten Ergebnis gekommen ($t(31) = -0,71$, $p = 0,48 > 0,05$).

Um zu überprüfen, ob die beiden Gruppen sich hinsichtlich ihrer Differenzen zu den korrekten Sätzen genauso verhalten, schauen wir auf die Ergebnisse der Post-hoc-Analyse der Differenz-Variablen im Anhang 9. Wie wir sehen, wurde beim paarweisen Vergleich der Gruppen G und H ein angepasstes Signifikanzniveau von $p = 1,000$ berechnet, was bedeutet, dass der Unterschied zwischen diesen Gruppen nicht signifikant ist. Wenn wir beide Differenz-Variablen direkt miteinander vergleichen und einen Wilcoxon-Test durchführen, bekommen wir wiederum ein Ergebnis, das auf einen unsignifikanten Unterschied hindeutet ($z = -1,563$, $p = 0,118 > 0,05$).

In Betracht der dargelegten Ergebnisse können wir an dieser Stelle definitiv davon ausgehen, dass der Bewertungsunterschied zwischen den Gruppen mit Genus- sowie Kasusfehlern unbedeutend ist und aus diesem Grund kann die Hypothese 2.1 nicht bestätigt werden.

4.2.3.2. Hypothese 2.2

In der Hypothese 2.2 war angenommen worden, dass der Verbmorphologie-Fehler (Gruppe J), der in dieser Arbeit als fehlerhafte Endung des Partizips II manifestiert wurde, von deutschen Muttersprachlern weniger toleriert wird als die falsche Verbposition im Haupt- (Gruppe K) bzw. im Nebensatz (Gruppe L). Statistisch betrachtet, können wir diese

Hypothese einerseits in zwei Unterhypothesen, nämlich $M(J) > M(K)$ und $M(J) > M(L)$ unterteilen. Andererseits können wir den Mittelwert der Gruppen K und L als eine neue Variable berechnen und sie mit der Mittelwertvariable der Gruppe J vergleichen: $M(J) > M(K+L)$.

Bevor wir mit der Untersuchung der Hypothesen anhand von Signifikanztests anfangen, beobachten wir die deskriptiven Daten dreier Gruppen. In unserer Analyse erwies die Gruppe J einen Mittelwert von ($M = 3,12$, $SD = 0,93$), der im Vergleich zu dem der Gruppen K ($M = 3,13$, $SD = 0,99$) und L ($M = 3,29$, $SD = 0,93$) sogar niedriger war. Betrachten wir die Medianposition dieser Gruppen im Boxplot-Diagramm (66), sehen wir, dass der Median der Gruppe J ($Med = 2,91$) zwar höher als der der Gruppe K ($Med = 2,87$), aber niedriger als die Medianlage der Gruppe L ($Med = 3,25$) ist. An dieser Stelle können wir definitiv davon ausgehen, dass die Gruppe J besser bewertet wurde als die Gruppe L.

Da es in dieser Hypothese um den Vergleich von mehr als zwei Variablen geht und die Mittelwert-Variable der Gruppe (K + L) in der bisher durchgeführten Varianzanalyse nicht berücksichtigt wurde, führen wir für die Überprüfung der Hypothese 2.2 eine neue Varianzanalyse bzw. einen Friedman-Test durch. Der Unterschied zwischen den vier Gruppen J, K, L und K+L wurde im Rahmen dieses Tests als unsignifikant ausgewiesen ($\chi^2(3) = 7,40$, $p = 0,06 > 0,05$), was darauf hindeutet, dass kein Gruppen-Paar sich bedeutend voneinander unterscheidet. Vergleichen wir die drei Fehlergruppen-Paare (J und K), (J und L) und (J und K+L) durch Wilcoxon-Tests, kommen wir für alle drei Vergleiche auch zu Ergebnissen, die auf unsignifikante Unterschiede hinweisen (s. Anhang 10.A).

Um zu überprüfen, ob die Lageunterschiede zwischen den Differenz-Variablen der vier Gruppen genauso unsignifikant sind, wenden wir das gleiche Verfahren für die Differenz-Gruppen an. Der Friedman-Test zeigte bei diesem Verfahren ein Ergebnis ($\chi^2(3) = 14,299$, $p = 0,003 < 0,01$), das auf einen signifikanten Unterschied zwischen mindestens zwei Paaren hindeutete. Um herauszufinden, welche Variablen sich signifikant voneinander unterscheiden, führten wir eine Post-hoc-Analyse durch, die darauf hinwies, dass nur die Differenz-Variablen der Gruppen K und L, deren Vergleich für die Hypothese 2.2 irrelevant ist, signifikant verschieden sind (s. Anhang 10.B). Wenn wir die Differenz-Variablen anhand eines Wilcoxon- oder t-Tests miteinander vergleichen, erzielen wir wiederum Ergebnisse, die auf nicht signifikante Unterschiede hinweisen (s. Anhang 10.A).

Laut dieser Beobachtungen können wir behaupten, dass die Fehler in der Realisierung des Partizips II nicht signifikant weniger toleriert werden als die falsche Positionierung des Verbs im Haupt- sowie im Nebensatz, und somit gilt die Hypothese 2.2 als nicht bestätigt.

4.2.3.3. Hypothese 2.3

Die Hypothese 2.3 basierte auf der Annahme, dass die fehlerhafte Verbposition im Hauptsatz (Gruppe K) von deutschen Muttersprachlern schlechter beurteilt wird als die falsche Position des Verbs im Nebensatz (Gruppe L). Allein die Betrachtung der deskriptiven Werte der beiden Gruppen, zeigt uns, dass die Gruppe K ($M = 3,13$, $SD = 0,99$, $Md = 2,87$) sogar besser bewertet wurde als die Gruppe L ($M = 3,29$, $SD = 0,93$, $Md = 2,87$).

Um zu sehen, ob die Gruppe K signifikant als weniger störend empfunden wurde, werfen wir einen Blick auf das Ergebnis der Post-hoc-Analyse für das Paar K und L in der Tabelle (69). Das angepasste Signifikanzniveau von $p = 1,00$ weist darauf hin, dass der Unterschied zwischen diesen zwei Gruppen unbedeutend ist. Vergleichen wir die beiden Gruppen anhand eines Wilcoxon-Tests, kann eine signifikante Differenz zwischen den beiden festgestellt werden ($z = -2,392$, $p = 0,017 < 0,05$, $r = -0,42$).

Die Post-hoc-Analyse der Differenz-Variablen im letzten Abschnitt (s. Anhang 10.B) zeigte uns, dass die Gruppe K mit einem angepassten Signifikanzniveau von $p = 0,002$ signifikant weniger von ihren parallelen korrekten Sätzen differierte als die Gruppe L. Betrachten wir jedoch die Ergebnisse der Post-hoc-Analyse der Differenz-Variablen im Anhang 9, finden wir für dieses Paar einen nicht signifikanten Wert von $p = 0,73$. Warum das gleiche Paar in zwei Post-hoc-Analysen unterschiedliche Signifikanzniveaus aufweist, lässt sich nur mit der Anzahl der Paar-Vergleiche in den jeweiligen Analysen begründen. Vergleichen wir die Differenz-Variablen der Gruppen K und L mithilfe eines Wilcoxon-Tests, stoßen wir auf Ergebnisse ($z = -2,968$, $p = 0,003 < 0,01$, $r = -0,52$), die auf signifikante Unterschiede hinweisen.

Auch wenn nicht bei allen Testverfahren der Unterschied zwischen den Gruppen K und L als signifikant gekennzeichnet wird, gehen wir angesichts der Tatsache, dass die Gruppe K sogar eine bessere Beurteilung als die Gruppe L aufweist, davon aus, dass auch die Hypothese 2.3 nicht als verifiziert gelten kann.

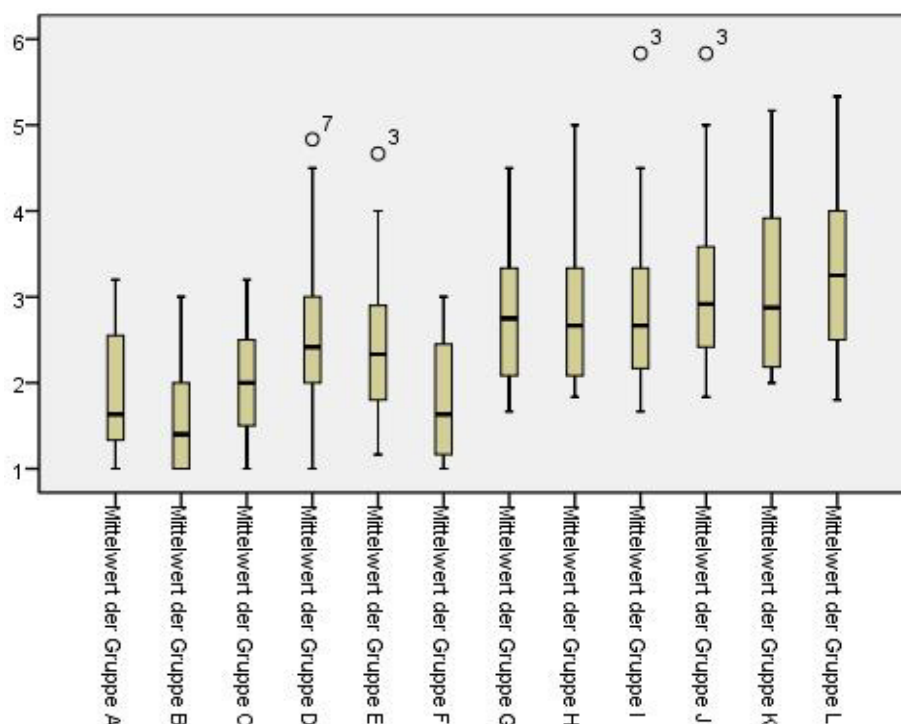
4.3. Überprüfung der Hypothesen zum kontrastiven Vergleich der Grammatik- und Aussprachefehler

4.3.1. Hypothese 3

Der Ausgangsgedanke unserer Arbeit bestand darin, dass Aussprachefehler die sprachliche Integration von Migranten mehr beeinträchtigen als grammatische Fehler. Basierend auf dieser Annahme formulierten wir die dritte Hypothese, in der wir davon ausgingen, dass phonologische Fehler in der Wahrnehmung der Muttersprachler störender wirken als syntaktische oder morphosyntaktische.

Um einen Überblick über die Bewertung der einzelnen Gruppen zu gewinnen, schauen wir zuerst auf die deskriptiven Daten im Anhang 8. Wie wir sehen, sind die höchsten Kennwerte der zentralen Tendenz, d. h. der Mittelwert und Median, in den phonologischen Fehlergruppen, die für die Gruppe D aufgezeichnet wurden, niedriger als der Mittelwert und Median aller grammatischen Fehlergruppen. Dies kann als ein Hinweis dafür dienen, dass die phonologischen Fehlerarten sogar als weniger störend bewertet wurden als die syntaktischen bzw. morphosyntaktischen. Mithilfe folgender Boxplots lassen sich die zentrale Tendenz und Variabilität der einzelnen Gruppen besser veranschaulichen.

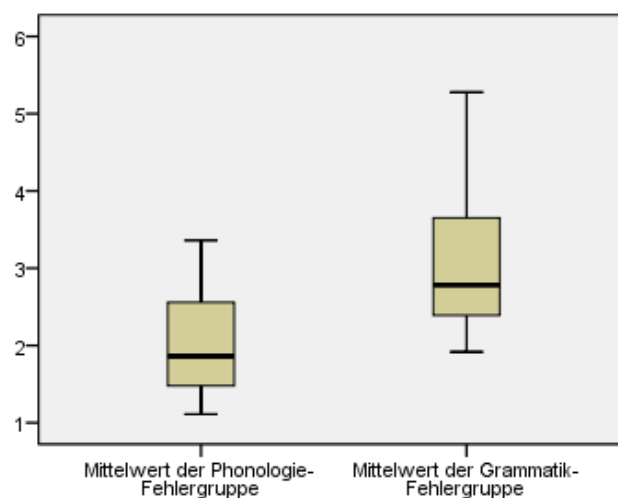
(70) Vergleichende Darstellung des Medians und der Streuung für die phonologischen und grammatischen Fehlerarten



Um zu überprüfen, ob die Aussprachefehler bedeutend weniger störend wirkten, berechneten wir für die beiden Gruppen eine neue Variable, die aus Mittelwerten der Bewertungen von sechs Untergruppen der jeweiligen Gruppe bestand. So ergaben sich für die Mittelwert-Variable der phonologischen Fehlergruppen die Kennwerte ($M = 2,04$, $SD = 0,65$) und für die grammatischen Fehlerarten die Maßzahlen ($M = 3,00$, $SD = 0,85$).

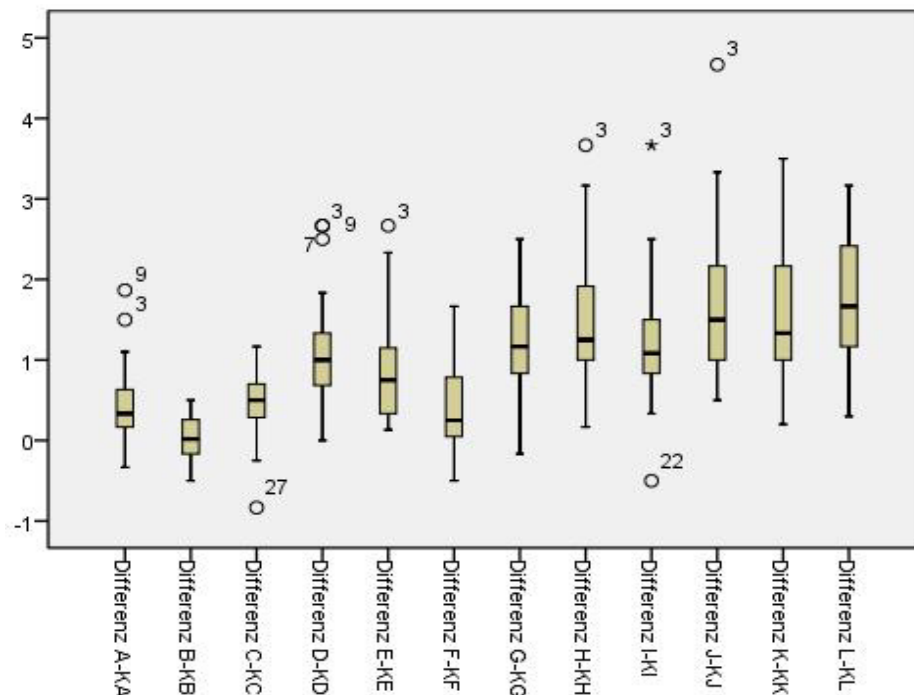
Der Mittelwertunterschied zeigte sich laut den Ergebnissen des t-Tests als signifikant ($t(31) = -10,582$, $p = 0,000 < 0,01$, $r = 0,88$), woraus wir schließen konnten, dass Grammatikfehler von den Probanden unserer Studie signifikant weniger toleriert wurden als Abweichungen in der Aussprache. Die unterschiedliche Lage des Medians, die in folgendem Boxplot-Diagramm visualisiert wird, bestätigt diese Annahme.

(71) Vergleichende Darstellung des Medians und der Streuung für die Mittelwert-Variablen der phonologischen und grammatischen Fehlerarten



Da in unserer Untersuchung für jeden Satz eine korrekte Version dargeboten wurde und die Bewertungsdifferenz zwischen dem fehlerhaften und dem korrekten Satz beim Vergleich der Fehlergruppen eine nicht zu unterschätzende Rolle spielt, schauten wir auch auf die Differenz-Variablen der phonologischen sowie grammatischen Fehlergruppen. Allein die Betrachtung der deskriptiven Werte für die Differenz-Variablen im Anhang 8 verdeutlicht, dass die Maßzahlen der zentralen Lage bei allen grammatischen Differenz-Variablen höher sind als bei den phonologischen. Die folgenden Boxplots visualisieren diese Annahme.

(72) Vergleichende Darstellung des Medians und der Streuung für die Differenz-Variablen der phonologischen und grammatischen Fehlerarten



Um zu sehen, ob die grammatischen Fehlergruppen signifikant größere Differenzen zu ihren parallelen korrekten Sätzen aufzeigen, bildeten wir auch hier zwei neue Variablen, die aus Mittelwerten der Untergruppen der phonologischen sowie grammatischen Differenz-Variablen bestanden. Die Differenz-Variable der Aussprache-Fehlergruppe zeigte bei dieser Analyse einen Mittelwert von ($M = 0,55$, $SD = 0,34$), der im Vergleich zu der syntaktischen bzw. morphosyntaktischen Differenz-Variable ($M = 1,51$ $SD = 0,67$) deutlich niedriger war. Da die beiden Variablen als normalverteilt erwiesen wurden, verglichen wir sie anhand eines t-Tests. Dieser Test ergab einen hochsignifikanten Unterschied zwischen den phonologischen sowie grammatischen Differenz-Variablen ($t(31) = -10,582$, $p = 0,000 < 0,01$, $r = 0,88$).

Wie bereits im Kapitel 3.3.1. erwähnt, wurde bei der Konzipierung des Untersuchungsmaterials dieser Studie darauf geachtet, dass jeder Satz in drei Versionen, nämlich einmal fehlerfrei, einmal mit einem Aussprache- und einmal mit einem Grammatikfehler, vorgespielt wird. Diese Gestaltungsform sollte uns ermöglichen, die 36 fehlerhaften Satzpaare, die jeweils aus zwei gleichen Sätzen, nämlich einmal mit einem grammatischen und einmal mit einem phonologischen Fehler, bestanden, kontrastiv miteinander zu vergleichen. Da in unserer Arbeit die meisten Sätze nicht normalverteilte Bewertungen aufwiesen und wir sie einer einheitlichen Analyse unterziehen wollten,

verglichen wir alle Satzpaare anhand des Wilcoxon-Tests. Wenn wir die Ergebnisse dieser Analyse im Anhang 11 betrachten, können wir feststellen, dass von 36 Paaren 33 auf eine signifikant schlechtere Bewertung des Grammatikfehlers hinweisen.

Angesichts der erzielten Ergebnisse in diesem Abschnitt lässt sich konstatieren, dass die Grammatikfehler in unserer Arbeit von den deutschen Probanden signifikant weniger akzeptiert wurden als die Abweichungen der Aussprache, und somit gilt die Hypothese 3 als nicht bestätigt.

4.3.2. Hypothese 4

In den ersten drei Hypothesen ging es um die Analyse der Bewertung verschiedener Fehlerarten. In der folgenden Hypothese befassen wir uns mit der Erkennung dieser Fehler durch die deutschen Muttersprachler und den Grad ihrer Sicherheit bei dieser Erkennung. Mit dieser Hypothese wurden also zwei Fragen aufgeworfen, die wir in zwei Unterhypothesen gliedern.

In der Hypothese 4.1 ist zu prüfen, ob die grammatischen Fehlerarten eine höhere Anzahl der korrekten Erkennung aufweisen als die phonologischen. In der Hypothese 4.2 versuchen wir die Frage zu beantworten, ob die Probanden bei der Erkennung der Grammatikfehler sicherer sind als bei der Wahrnehmung der Aussprachefehler.

Bevor wir mit der Überprüfung der Hypothesen anfangen, werfen wir erneut einen Blick auf die für diese Hypothese vorgesehenen Fragen in unserem Fragebogen im Anhang 3. Wie wir sehen, sollten die Probanden nach der Bewertung des Satzes für jeden Satz, falls sie ihn als fehlerhaft beurteilt hatten, aufschreiben, welchen Fehler sie gehört haben. Für die Datenerhebung zur Erkennung des Fehlers durch die Probanden wurden die Fehlerangaben in die fünf folgenden Antwortkategorien eingeteilt:

- 0 = keine Angabe
- 1 = Fehler richtig erkannt
- 2 = fälschlicherweise Aussprachefehler gehört
- 3 = fälschlicherweise Grammatikfehler gehört
- 4 = fälschlicherweise Stil-/Wortschatz- oder andere Fehler gehört
- 5 = als unverständlich beurteilt

Die Zuordnung der Zahlen 0 bis 5 hat, wie wir wissen, keine empirische Bedeutung und präsentiert nur die verschiedenen Ausprägungen der Antworten (vgl. Raab-Steiner & Benesch

2008:24). Da wir hier mit der Auswertung der nominalskalierten Daten zu tun haben, müssen wir uns für die Überprüfung der Hypothese 4.1 auf Verfahren beschränken, die für nominale Skalen geeignet sind, nämlich auf die Analyse der Häufigkeitsverteilungen.

Für die Erhebung der Daten zum Sicherheitsgrad bei der Erkennung des Fehlers sollten die Probanden in einer Vier-Punkte-Skala ankreuzen, wie sicher sie sich bei dieser Erkennung waren. Wie im Abschnitt 3.3.4. erwähnt, sollte uns die Verwendung der endpunktbenannten Skalen ermöglichen, auch bei der Überprüfung der Hypothese 4.2 die Auswertungsmöglichkeiten der intervallskalierten Daten anzuwenden.

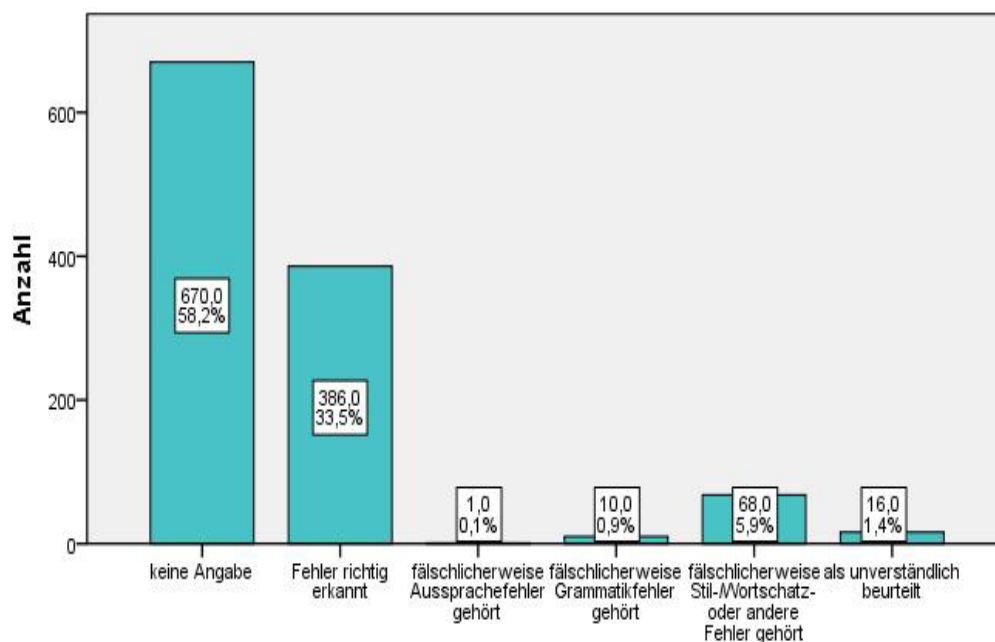
4.3.2.1. Hypothese 4.1

Zur Überprüfung der Hypothese 4.1, in der angenommen wird, dass die Anzahl der korrekten Erkennung der Grammatikfehler höher ist als die der Aussprachefehler, stellen wir zuerst anhand folgender Balkendiagramme die Häufigkeitsverteilung der verschiedenen Antwortkategorien beider Fehlergruppen dar. Wie die Abbildungen (73) und (74) zeigen, gibt es für die Bestätigung der Hypothese 4.1 augenfällige Indizien. Wie erwartet, wurden bei der Grammatik-Fehlergruppe 82% der Fehler auch als solche erkannt, während diese Zahl bei der Gruppe der Aussprachefehler 33,5% betrug. Der prozentuale Unterschied von ca. 49% wurde anhand des McNemar-Tests³² auf Signifikanz überprüft und dabei wurde ein hochsignifikanter Unterschied zwischen den Häufigkeiten der korrekten Erkennung in beiden Fehlergruppen bewiesen ($\chi^2 (1, N=1147)^{33} = 503,47, p = 0,000 < 0,01, w = 0,66^{34}$). Die Kreuztabelle der Fehlerangaben beider Gruppen ist im Anhang 12.A einzusehen.

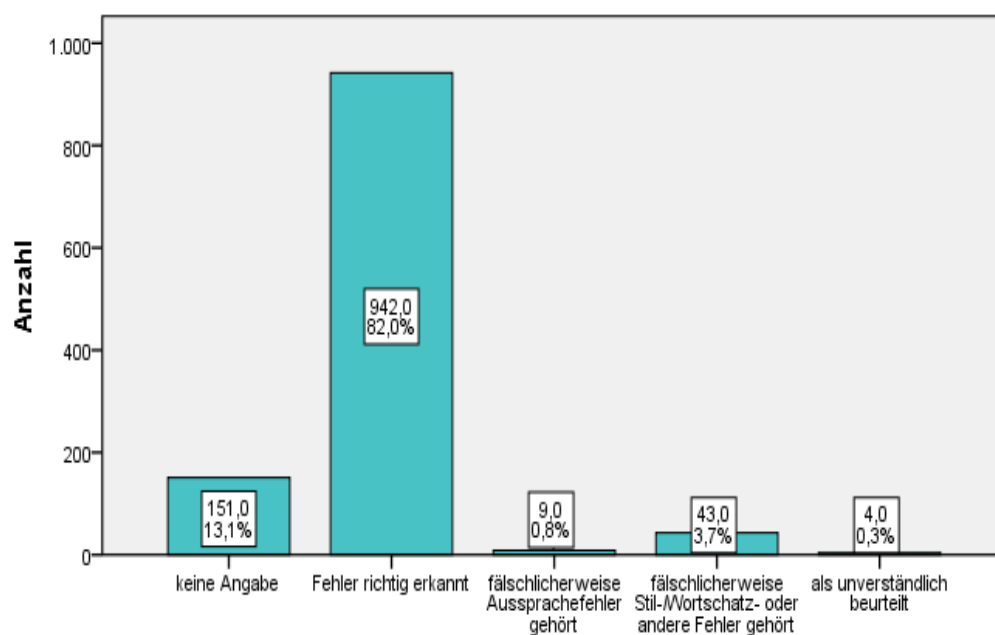
³² Dieser Test ist ein Prüfverfahren für nominalskalierte Daten und wird eingesetzt, wenn dieselbe Stichprobe zweimal auf ein alternatives Merkmal hin untersucht wird. Der McNemar-Test setzt voraus, dass die beiden verbundenen Stichproben gleich groß und die zu untersuchenden Variablen dichotom bzw. zweifach gestuft sind (vgl. Bortz & Schuster 2010:143). Da uns bei der Überprüfung der Hypothese 4.1 nur der Häufigkeitsunterschied zwischen der korrekten und der falschen Erkennung interessierte und um die Voraussetzungen des McNemar-Tests zu erfüllen, reduzierten wir die fünffach gestuften Antwortkategorien in dieser Arbeit auf zwei Stufen, nämlich auf die korrekte und die falsche Erkennung des Fehlers. So wurden für die Durchführung des McNemar-Tests alle Antwortkategorien außer „Fehler richtig erkannt“ in eine Kategorie, nämlich als „Fehler nicht erkannt“ zusammengefasst.

³³ Da SPSS für den McNemar-Test nur das Signifikanzniveau ausgibt, wurde die Prüfgröße anhand folgender Formel berechnet: $\chi^2 = (b - c)^2 / b + c$. Da es sich um eine 2*2-Tabelle handelte, wurde als Freiheitsgrad die Zahl 1 angegeben und die Anzahl von N präsentierte die Stichprobengröße (vgl. Bortz&Schuster2010:147).

(73) Häufigkeitsverteilung der Fehlerangaben für die Gruppe der Aussprachefehlerarten



(74) Häufigkeitsverteilung der Fehlerangaben für die Gruppe der Grammatikfehlerarten



Bei der Formulierung der Hypothese 4 im Abschnitt 3.1. wurde außerdem vermutet, dass die Abweichungen der Aussprache mehr als die Grammatikfehler zur fälschlichen Erkennung des Fehlers führen. Wenn man sich die oben dargestellten Balkendiagramme ansieht, stellt man

³⁴ Die Effektgröße w wurde mithilfe der Formel $w = \sqrt{\chi^2 / N}$ berechnet. Diese Formel wird für alle χ^2 -Tests verwendet, wobei N sich auf die Größe der Gesamtstichprobe bezieht (vgl. Schäfer 2011:154).

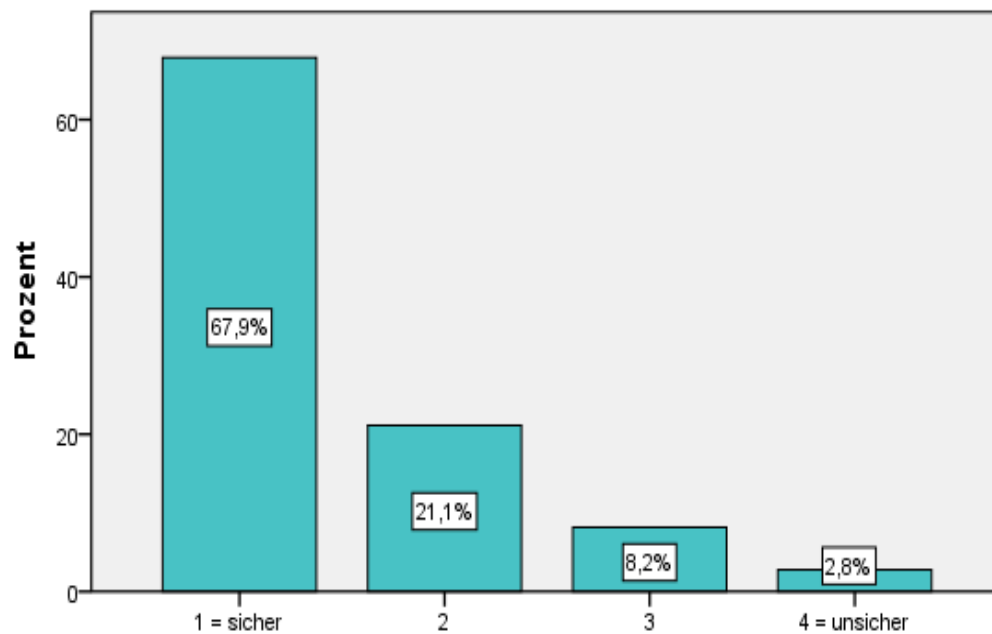
fest, dass die Prozentzahl der Antwortkategorien 2 bis 5, die sich auf die unterschiedlich falsche Erkennung der Abweichungen beziehen, bei den beiden Gruppen keine großen Unterschiede aufzeigt.

Weiterhin fällt auf, dass die Anzahl der fehlerhaften Sätze, bei denen keine Angaben zu den erkannten Fehlern gemacht wurden, in der Gruppe der Aussprachefehler mit 58,2% eindeutig höher ist als in der Grammatik-Fehlergruppe mit nur ca. 13%. Dieser Unterschied und die Tatsache, dass die Aussprachefehler in unserer Studie signifikant besser bewertet wurden, können darauf hinweisen, dass die in dieser Arbeit untersuchten Ausspracheabweichungen von den deutschen Probanden deutlich weniger wahrgenommen wurden.

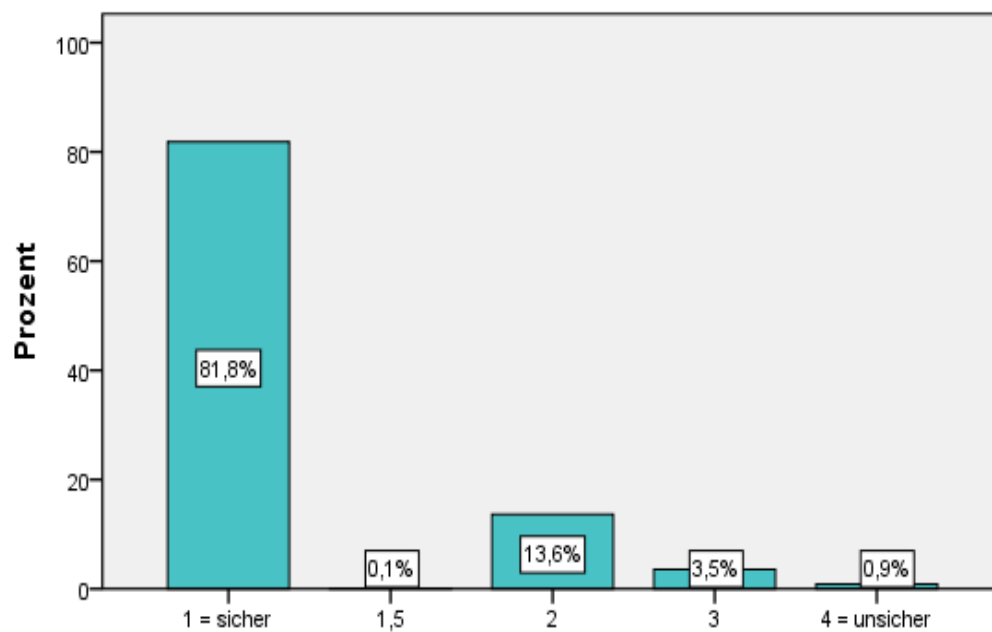
4.3.2.2. Hypothese 4.2

In der Hypothese 4.2 wurde angenommen, dass der Sicherheitsgrad bei der Erkennung der Grammatikfehler höher ist als bei der der Aussprachefehler. Für die Überprüfung dieser Hypothese wurden, wie bereits erwähnt, nur Angaben bzw. Sicherheitsgrade berücksichtigt, die bei einer korrekten Erkennung des Fehlers gemacht wurden. Wenn wir die Häufigkeit verschiedener Sicherheitsgrade, die für die Grammatik- sowie die Aussprache-Fehlergruppe berechnet wurden, anhand folgender Balkendiagramme betrachten, können wir feststellen, dass trotz eines minimalen Unterschieds die Anzahl der mit 1 angekreuzten Sicherheitsgrade bei der Gruppe der Grammatikfehler höher ist als bei der Gruppe der Ausspracheabweichungen.

(75) Häufigkeitsverteilung der Sicherheitsangaben bei der Erkennung der Aussprachefehler



(76) Häufigkeitsverteilung der Sicherheitsangaben bei der Erkennung der Grammatikfehler



Um zu überprüfen, ob die angegebenen Sicherheitsgrade in beiden Gruppen bedeutend verschieden sind, berechneten wir zuerst den Mittelwert des Sicherheitsgrads in beiden Gruppen. Dabei zeigten die Sicherheitsgrade der Aussprache-Fehlergruppe einen Mittelwert von ($M = 1,45$, $SD = 0,76$), der leicht höher war als bei der Grammatik-Fehlergruppe ($M =$

1,23, $SD = 0,55$).³⁵ Anhand eines Wilcoxon-Tests konnten wir feststellen, dass die Grammatikfehler signifikant sicherer bewertet wurden als die Aussprachefehler ($z = -5,557$, $p = 0,000 < 0,01$, $r = -0,16$). Da der berechnete r -Wert bzw. die Effektgröße jedoch recht niedrig ist, bleibt umstritten, ob wir die Hypothese 4.2 als bestätigt betrachten können.

Wenn wir nur die unterschiedliche Anzahl der mit 1 angekreuzten Sicherheitsgrade in beiden Gruppen anhand eines McNemar-Tests, dessen Kreuztabelle im Anhang 12.B zu finden ist, miteinander vergleichen³⁶, ergibt sich ein signifikanter Unterschied zwischen den beiden Gruppen ($\chi^2(1, N=1146) = 200$, $p = 0,000 < 0,01$, $w = 0,41$), wobei die Effektgröße hier einen nahezu mittleren Effekt aufweist.

³⁵ Die niedrigere Zahl verweist hier auf einen höheren Sicherheitsgrad, da in der Skala für den Sicherheitsgrad die Werte 1 als „sehr sicher“ und 4 als „sehr unsicher“ definiert wurden.

³⁶ Dafür wurden die Sicherheitswerte 2 bis 4 unter einen Wert, nämlich 0 zusammengefasst. So konnten wir zwei Kategorien, nämlich 1(sicher) und 0 (unsicher), mithilfe einer Kreuztabelle und eines McNemar-Tests miteinander vergleichen.

4.4. Überprüfung der Hypothesen zur allgemeinen Wahrnehmung der Sprechfehler

4.4.1. Hypothese 5

Die fünfte Hypothese warf die Frage auf, ob es einen direkten Zusammenhang zwischen der korrekten Erkennung des Fehlers und dessen Bewertung gibt. Wie bereits in den Hypothesen 3 und 4.1 festgestellt, erhielten die Grammatikfehler im Durchschnitt signifikant schlechtere Bewertungen und wiesen eine höhere Anzahl der korrekten Erkennung des Fehlers auf als die Aussprachefehler. Auch wenn wir die Fehlerangaben der einzelnen Gruppen in den Kapiteln 4.1.1. und 4.2.1. betrachten, können wir feststellen, dass sowohl bei den Grammatik- als auch bei den Aussprache-Fehlergruppen Fehlerarten, die in den Bewertungen der Probanden schlechter abschnitten, z. B. die Aussprache-Fehlergruppen D und E und die Grammatik-Fehlerarten J und L, gleichzeitig am häufigsten als Fehler erkannt wurden. Die Gruppe B, die besonders positiv bewertet worden war, wies z. B. nur 1,6% der korrekten Erkennung auf und die Gruppe G, die unter den Grammatikfehlerarten als weniger störend beurteilt wurde, zeigte für die korrekte Erkennung eine Zahl von 67%, die im Vergleich zu den anderen Grammatikfehlern niedriger war.

In Anbetracht dieser Ergebnisse kann man davon ausgehen, dass je mehr Fehler in einer Gruppe erkannt wurden, desto schlechter sich die Bewertung dieser Gruppe zeigte. Um zu überprüfen, inwieweit diese schlechte Bewertung durch die korrekte Erkennung zu erklären ist, soll zuerst die Korrelation zwischen den Variablen „Fehlererkennung“ und „Bewertung“ ausgerechnet werden. Da bei der Berechnung des Zusammenhangs die höhere Anzahl der Fehlererkennung der schlechteren Bewertung entsprechen sollte und in unserer Arbeit eine schlechte Bewertung mit einem höheren Skalenniveau definiert ist, muss die Korrelation dabei positiv sein.

Die Variable „Fehlererkennung“ wurde auch hier als eine dichotome Variable, in der zwei Ausprägungen der Fehlererkennung, nämlich „Fehler erkannt“ und „Fehler nicht erkannt“, mit den Werten 1 und 0 codiert wurden, präsentiert. Die Korrelation zwischen dieser Variable

und der Variable „Bewertung“, die wir als intervallskaliert annehmen, wurde durch Pearsons Produkt-Moment-Korrelation³⁷ berechnet.

Der Zusammenhang zwischen den zwei Variablen³⁸ erwies sich dabei als positiv und mittelschwach³⁹ ($r(2298^{40}) = 0,44$), war allerdings hochsignifikant ($p = 0,000 < 0,01$). Wenn wir die Mittelwerte der zwei Ausprägungen „Fehler nicht erkannt“ und „Fehler erkannt“ in der unteren Tabelle betrachten, können wir die oben berechnete Korrelation folgendermaßen interpretieren: Es besteht eine Korrelation zwischen der Erkennung des Fehlers und der Bewertung in der Form, dass die korrekt erkannten Fehler negativer bewertet wurden als die nicht erkannten, obwohl dieser Zusammenhang relativ schwach ist.

³⁷ Zur Berechnung der Korrelation zwischen einem dichotomen Merkmal und einem intervallskalierten Merkmal wird die punktbiseriale Korrelation herangezogen. Wenn die zwei Ausprägungen der dichotomen Variable durch die Werte 0 und 1 ersetzt werden, ist das Ergebnis der punktbiserialen Korrelation identisch mit dem Ergebnis von Pearsons Produkt-Moment-Korrelation. Obwohl die Pearson-Korrelation eine Normalverteilung der Daten voraussetzt, erweist sich deren Signifikanztest als robust gegenüber Verletzungen dieser Voraussetzung (vgl. Bortz & Schuster 2010: 162-171).

³⁸ Die zwei Variablen präsentierten die Bewertungen und die Angaben zu der Fehlererkennung von 32 Probanden für 12 Fehlerarten, die jeweils aus sechs Sätzen bestanden, d. h. jede Variable bestand aus 2304 Fällen, wobei einige davon als fehlende Werte eliminiert wurden.

³⁹ Für die Interpretation der Stärke der Korrelationskoeffizienten greifen wir auf die von Albert und Marx (2014: 130) präsentierte allgemeine Faustregel zurück:

Wert	Interpretation
0,90-1	sehr hohe Korrelation, sehr starke Beziehung
0,70-0,89	hohe Korrelation, ausgeprägte Beziehung
0,40-0,69	mittlere Korrelation, substantielle Beziehung
0,20-0,39	schwache Korrelation, definitive, aber geringe Beziehung
0-0,19	keine bis leichte Korrelation

⁴⁰ Der Freiheitsgrad (df) wurde durch die Formel $(n-2)$ berechnet, wobei (n) die Anzahl der gepaarten Vergleiche und 2 die Anzahl der Gruppen präsentierte.

(77) Mittelwertvergleich der zwei Ausprägungen der Fehlererkennung

Fehlererkennung	Mittelwert	Häufigkeit	Standardabweichung	Median
Fehler nicht erkannt (0)	1,873	972 (42,3%)	1,0612	2,000
Fehler erkannt (1)	2,934	1328 (57,7%)	1,0815	3,000
Gesamtsumme	2,486	2300 ⁴¹	1,1938	2,000

Wenn wir das Quadrat von dem oben aufgeführten Korrelationskoeffizienten r berechnen,⁴² stellen wir fest, dass nur 19,3% der negativen Bewertung durch die korrekte Erkennung des Fehlers erklärt wurden. Dies bedeutet, dass die Erkennung des Fehlers für dessen negative Bewertung zwar relevant ist, es aber andere Faktoren gibt, die die weiteren 80,7% der schlechten Bewertung erklären.

Da in unserer Arbeit die Aussprache- und Grammatikfehler laut den Ergebnissen der Hypothese 1 sehr unterschiedlich bewertet wurden, wäre an dieser Stelle interessant, zu überprüfen, wie die Korrelation zwischen der Bewertung und der Fehlererkennung in den Aussprache- und Grammatik-Fehlergruppen aussieht. Die folgende Tabelle stellt die Ergebnisse der Korrelationsanalyse in diesen zwei Gruppen dar. Der Vollständigkeit halber wurden gleich Mittelwerte und Häufigkeiten der Ausprägungen „Fehler erkannt“ und „Fehler nicht erkannt“ präsentiert.

⁴¹ Die fehlenden Werte der Fehlerangaben, die sich auf irrelevante Angaben wie Rechtschreibfehler bezogen, wurden aus der Analyse herausgenommen. Die fehlenden Werte der Bewertung, bei denen der falsche Satz besser beurteilt wurde als die korrekte Version, wurden dagegen in der Analyse beibehalten, weil es sich hier nur um den Zusammenhang zwischen der Bewertung fehlerhafter Sätze und der korrekten Erkennung des Fehlers handelte.

⁴² Das r -Quadrat, auch Determinationskoeffizient r^2 genannt, zeigt uns, wie viel von der Varianz einer Variable durch die Varianz einer anderen Variable erklärt werden kann. Um zu sehen, wie viel Prozent der Streuung in einer Variable von der Streuung der anderen Variable beeinflusst wird, multipliziert man die Anzahl der r^2 mit 100 (vgl. Albert & Marx 2014: 159).

(78) Korrelation zwischen der Bewertung und Fehlererkennung in den Grammatik- sowie Aussprache-Fehlergruppen

	Pearson-Korrelation	% r^2		Mittelwert	Häufigkeit
Aussprache-Fehler (N=1151)	$r(1149) = \mathbf{0,423^{43}}$ Sig.(2-seitig) = 0,000	17,89%	Fehler erkannt (1)	2,58	386 (33,5%)
			Fehler nicht erkannt (0)	1,68	765 (66,5%)
Grammatik-Fehler (N=1149)	$r(1147) = 0,168$ Sig.(2-seitig) = 0,000	2,82%	Fehler erkannt (1)	3,07	942 (82%)
			Fehler nicht erkannt (0)	2,57	207 (18%)

Wie man sieht, ist die Korrelation in der Aussprache-Fehlergruppe etwas stärker ausgeprägt als in der Grammatik-Fehlergruppe. Die Tatsache, dass die Gruppe der Grammatikfehler trotz einer schlechteren Bewertung und höheren Anzahl der erkannten Fehler eine niedrigere Korrelation aufweist, kann damit begründet werden, dass in dieser Gruppe auch die nicht erkannten Fehler als fast genauso störend bewertet wurden wie die erkannten.. In der Aussprache-Fehlergruppe ist die Mittelwert-Differenz zwischen den erkannten und nicht erkannten Fehlern etwas größer und weist deshalb eine stärkere Korrelation auf.

Es wäre anschließend interessant, diese Untersuchung etwas zu konkretisieren und zu überprüfen, wie unterschiedlich die Korrelation in den Untergruppen der Grammatik- und Aussprache-Fehlergruppen aussieht. Die untere Tabelle zeigt die Korrelation zusammen mit den Mittelwerten und Häufigkeiten der zwei Ausprägungen der Fehlererkennung in 12 Untergruppen. Wie man sieht, sind die Korrelationen der Gruppen F (falsche Akzentposition) und A (falsche Realisierung der Affrikate [ts]) eindeutig stärker ausgeprägt als bei allen anderen Gruppen. Die Gruppe D (falsche Realisierung des Umlauts), die bei der Bewertung der Aussprachefehler am schlechtesten abschnitt, und die Grammatik-Fehlergruppen J (Präpositionsfehler), K (falsche Verbposition im Hauptsatz) und L (falsche Verbposition im Nebensatz), die von den Probanden weniger als die anderen Grammatikfehler toleriert wurden, weisen eine extrem schwache Korrelation auf, wobei diese bei den letzten drei Gruppen sogar negativ ist.

⁴³ Die mittleren bis hohen Korrelationen sind fett markiert.

Wenn wir die Mittelwert-Differenz zwischen den zwei Ausprägungen der Fehlererkennung bei den Gruppen mit schwacher Korrelation bzw. schlechterer Bewertung anschauen, können wir feststellen, dass die nicht erkannten Fehler genauso schlecht oder bei den negativen Korrelationen sogar schlechter bewertet wurden als die erkannten Fehler. Möglicherweise war die störende Wirkung dieser Fehler so stark, dass die Probanden sie auch im Falle der Nichterkennung unbewusst als sehr negativ bewertet haben. Die weniger störenden Aussprachefehler zeigten dagegen eine relativ höhere Korrelation, weil bei diesen Gruppen die negative Bewertung einen höheren Abhängigkeitsgrad von der Erkennung des Fehlers aufwies.

Aufgrund der oben festgestellten Ergebnisse können wir davon ausgehen, dass die korrekte Erkennung des Fehlers dessen negative Bewertung zwar beeinflusst, es aber, wie wir bei den Korrelationsergebnissen der Untergruppen gesehen haben, viele andere Faktoren, wie den allgemeinen Eindruck bzw. die unbewusste Wahrnehmung des Fehlers oder den Klang der fehlerhaften Sätze, gibt, die für die schlechte Bewertung verantwortlich sind.

(79) Korrelation zwischen der Bewertung und Fehlererkennung in den Untergruppen der Grammatik- und Aussprache-Fehlergruppen

N=188 df=186	Pearson-Korrelation	r^2 %		Mittelwert	Häufigkeit
Gruppe A	$r = \mathbf{0,447}$ Sig. (2-seitig) = 0,000	19,98%	Fehler erkannt (1)	2,431	51 (27,1%)
			Fehler nicht erkannt (0)	1,562	137 (72,9%)
Gruppe B	$r = 0,246$ Sig. (2-seitig) = 0,001	6,05%	Fehler erkannt (1)	3,000	3 (1,6%)
			Fehler nicht erkannt (0)	1,476	185 (98,4%)
Gruppe C	$r = 0,309$ Sig. (2-seitig) = 0,000	9,54%	Fehler erkannt (1)	2,564	39 (20,7%)
			Fehler nicht erkannt (0)	1,832	149 (79,3%)
Gruppe D	$r = 0,076$ Sig. (2-seitig) = 0,300	0,57%	Fehler erkannt (1)	2,599	162 (86,2%)
			Fehler nicht erkannt (0)	2,385	26 (13,8%)
Gruppe E	$r = 0,309$ Sig. (2-seitig) = 0,000	9,54%	Fehler erkannt (1)	2,756	78 (41,5%)
			Fehler nicht erkannt (0)	2,045	110 (58,5%)

Gruppe F	$r = \mathbf{0,512}$	26,21%	Fehler erkannt (1)	2,458	48 (25,5%)
	Sig. (2-seitig) = 0,000		Fehler nicht erkannt (0)	1,457	140 (74,5%)
Gruppe G	$r = 0,284$	8,06%	Fehler erkannt (1)	3,000	125 (66,5%)
	Sig. (2-seitig) = 0,000		Fehler nicht erkannt (0)	2,302	63 (33,5%)
Gruppe H	$r = 0,313$	9,79%	Fehler erkannt (1)	2,987	156 (83%)
	Sig. (2-seitig) = 0,000		Fehler nicht erkannt (0)	2,031	32 (17%)
Gruppe I	$r = 0,394$	15,52%	Fehler erkannt (1)	3,087	150 (79,8%)
	Sig. (2-seitig) = 0,000		Fehler nicht erkannt (0)	1,868	38 (20,2%)
Gruppe J	$r = -0,119$	1,41%	Fehler erkannt (1)	3,053	170 (90,4%)
	Sig. (2-seitig) = 0,104		Fehler nicht erkannt (0)	3,500	18 (9,6%)
Gruppe K	$r = -0,036$	0,12%	Fehler erkannt (1)	3,106	161 (85,6%)
	Sig. (2-seitig) = ,623		Fehler nicht erkannt (0)	3,222	27 (14,4%)
Gruppe L	$r = -0,126$	1,58%	Fehler erkannt (1)	3,233	163 (86,7%)
	Sig. (2-seitig) = 0,084		Fehler nicht erkannt (0)	3,640	25 (13,3%)

4.4.2. Hypothese 6

In der letzten Hypothese wurde angenommen, dass Sprecher mit ausländischem Akzent selbst bei fehlerfreien Sätzen den Eindruck entstehen lassen, sie machen einen Fehler. Statistisch ausgedrückt sollten die gesprochenen korrekten Sätze bei der Bewertung Werte erzielen, deren Durchschnitt größer ist als 1. Wenn wir die durchschnittliche Beurteilung der 36 korrekten Sätze berechnen, ergeben sich folgende Maßzahlen: ($M = 1,49$, $SD = 0,45$, $Md = 1,37$). Betrachten wir die deskriptiven Werte der einzelnen korrekten Sätze im Anhang 6, fällt uns auf, dass keiner der Sätze einen Mittelwert aufweist, der exakt bei 1 liegt. Bei einigen Sätzen wie bei den Sätzen 9, 15 und 25 lag die durchschnittliche Bewertung sogar über 2. Dies bedeutet, dass kein Satz von allen Hörern mit 1 bewertet wurde.

Da die Bewertung in der sechsstufigen Skala in dieser Arbeit die Antwort auf die Frage „Wie gut wurde der Satz auf Standarddeutsch gesprochen?“ reflektiert, können wir die angekreuzten Bewertungen nicht nur auf den Eindruck der Fehlerhaftigkeit des Satzes zurückführen. Bei der Abgabe ihrer Bewertung haben die Probanden offenbar nicht nur darauf

geachtet, ob der Sprecher Fehler macht, sondern es spielten auch andere Faktoren wie die Stimmwirkung, der allgemeine Eindruck oder sogar der Inhalt des Satzes bei ihrer Beurteilung eine Rolle. Um zu sehen, ob und inwiefern die korrekten Sätze als fehlerhaft empfunden wurden, analysieren wir die Fehlerangaben zu den korrekten Sätzen.

Wenn wir die Balkendiagramme, die die Fehlerangaben der 32 Probanden für jeden korrekten Satz präsentieren, im Anhang 13 betrachten, können wir feststellen, dass nur bei 6 von 36 Sätzen keine Angaben zu den erkannten Fehlern gemacht wurden, was wir hier als Einstufung des Satzes als fehlerfrei interpretieren.⁴⁴ Die folgenden vier Sätze, die bei der Bewertung besonders schlecht abschnitten, wiesen eine Prozentzahl von mehr als 30% der als Fehler angemerkten Angaben auf.

9. Er beteiligt sich immer aktiv am Unterricht.

15. Mein Zahn ist sehr empfindlich und tut beim Kauen weh.

25. Deutsch ist für mich eine zweite Sprache.

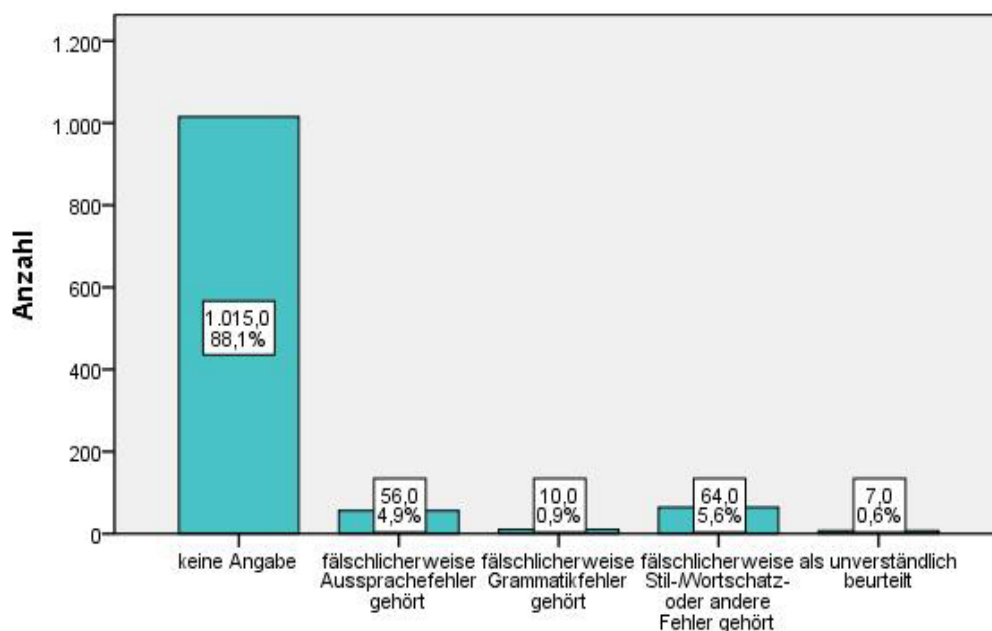
30. Obwohl er nicht so groß ist, kann er drei Meter hoch springen.

Bei der Bewertung des Satzes 9 kritisierten ca. 18% der Hörer die falsche Aussprache und ca. 15% sahen das Verb *beteiligen* als einen möglichen Wortschatz- bzw. Grammatikfehler an. Im Satz 15 wurde das Wort *Kauen* sowohl als Aussprachefehler (25%) als auch als eine unverständliche oder falsche Äußerung (15%) bezeichnet. Die Nominalphrase *eine zweite Sprache* wurde im Satz 25 von ca. 34% der Probanden als ein Stil- bzw. Wortschatzfehler gekennzeichnet, wobei alle Hörer als korrekte Form das zusammengesetzte Wort *Zweitsprache* vorgeschlagen haben. Im Satz 30 haben ca. 28% die falsche Aussprache bzw. die Betonung des Verbs *springen* als Fehler angegeben. Wenn wir die Balkendiagramme und dazu angegebenen Anmerkungen der Probanden im Anhang 13 betrachten, können wir feststellen, dass die meisten Anmerkungen sich auf die Aussprache- bzw. Stil- oder Wortschatzfehler bezogen haben.

⁴⁴ Bei der Durchführung des McNemar-Tests für die Hypothese 4.1 und für die Berechnung der Korrelationen in der Hypothese 5 wurden alle Antwortkategorien außer „Fehler richtig erkannt“, d. h. auch „keine Angaben“, in eine Kategorie, nämlich als „Fehler nicht erkannt“, zusammengefasst. Bei der Überprüfung dieser Hypothese, bei der es sich um die Untersuchung der korrekten Sätze handelt, gehen wir davon aus, dass „keine Angaben“ ein Indiz für die Fehlerlosigkeit des Satzes ist.

Die Abbildung (80), die die Fehlerangaben von 32 Probanden für 36 korrekte Sätze veranschaulicht, bestätigt diese Annahme. Wie man sieht, wurden die korrekten Sätze zu einem großen Teil, nämlich ca. 88% als fehlerfrei bewertet. In 5,6% der Sätze wurden Stil- bzw. Wortschatzfehler gehört und ca. 5% der Anmerkungen bezogen sich auf fälschlicherweise entdeckte Aussprachefehler. Weniger als 1% betrug die Angaben zu der fehlerhaften Erkennung der Grammatik- und Verständlichkeitsfehler. Insgesamt wurden also 12% der korrekten Sätze als fehlerhaft bezeichnet.

(80) Häufigkeitsverteilung der Fehlerangaben für die korrekten Sätze



Angeichts der Tatsache, dass der Sprecher in dieser Studie ein Nichtmuttersprachler war und beim Vorlesen der sowohl korrekten als auch fehlerhaften Sätze einen kontinuierlichen Akzent beibehielt, können wir die 12% der fälschlicherweise als fehlerhaft bewerteten Sätze auf seinen Akzent zurückführen und als einen Hinweis dafür betrachten, dass einige Hörer bei der Bewertung der Äußerungen eines Akzentsprechers neben den eingeplanten Fehlern andere Abweichungen in ihre Bewertung einschließen. Obwohl die 12-prozentige Anzahl darauf verweist, dass die Hypothese 6 nicht als falsifiziert betrachtet werden kann, wären für ihre Verifizierung noch andere Untersuchungen nötig. Man hätte z. B. einen Muttersprachler die korrekten Sätze sprechen lassen und den Unterschied zwischen der Bewertung seiner gesprochenen Sätze und der von unserem Sprecher vorgelesenen Sätze durch einen Signifikanztest vergleichen müssen. Die Durchführung einer solchen Untersuchung sprengt

allerdings den Rahmen dieser Arbeit und bleibt einer weiteren empirischen Forschung überlassen.

5. Zusammenfassung und Diskussion der Ergebnisse

Die statistische Analyse der gewonnenen Daten im Kapitel 4 lieferte uns neben den für die Überprüfung der Hypothesen relevanten Ergebnissen Informationen, die wir für die Klassifikation und Gewichtung der Fehlerarten verwenden konnten. Mithilfe der durchgeführten Varianz- und Folgeanalysen erstellten wir für die Aussprache- und Grammatik-Fehlergruppen eine Rangliste, die die Fehlerarten im Hinblick auf ihre störende Wirkung auf die deutschen Muttersprachler klassifizierte.

(81) Rangliste der Aussprache- und Grammatik-Fehlerarten

Platz	Aussprache-Fehlerarten	Grammatik-Fehlerarten
1	Gruppe D (Realisierung des Umlauts)	Gruppe L (Verbposition im Nebensatz)
2	Gruppe E (Konsonantenhäufung in der Wortanlautsilbe)	Gruppe J (Partizip II)
3	Gruppe C (Vokalquantität)	Gruppe K (Verbposition im Hauptsatz)
4	Gruppe A (Realisierung der Affrikate [ts])	Gruppe I (Präposition)
5	Gruppe F (Wortakzent)	Gruppe H (Kasus)
6	Gruppe B (Auslautverhärtung)	Gruppe G (Genus)

Da bei dieser Rangordnung sich einige Gruppen nur minimal voneinander unterschieden, bezogen wir uns für die Schlussfolgerung nur auf signifikante Unterschiede. In der Gruppe der Aussprachefehler wurden die falsche Aussprache des Umlauts und der Konsonantenhäufung in der Wortanlautsilbe besonders bzw. signifikant negativer evaluiert, wobei diese zwei Fehlerarten mehr als die anderen als Fehler identifiziert wurden. Die Verletzung der Auslautverhärtungsregel wurde dagegen von fast keinem Hörer als Fehler erkannt und schnitt in der Bewertung besonders positiv ab. In der Gruppe der Grammatikfehler wurden die drei Fehlerarten Verbposition im Haupt- und Nebensatz sowie Partizip II im Vergleich zu den anderen drei Gruppen signifikant negativer sanktioniert und zeigten gleichzeitig eine höhere Anzahl der korrekten Erkennung auf. Vergleichen wir unsere Ergebnisse mit den Resultaten anderer im Forschungsüberblick vorgestellten Studien, können wir nur im Bereich der Wirkung von Grammatik-Abweichungen eine Übereinstimmung mit den Ergebnissen von Politzer (1978) feststellen. In beiden Untersuchungen wurden die Abweichungen im Bereich der Verbmorphologie und Wortstellung im Satz signifikant schlechter beurteilt als Genus- und Kasusfehler. Die abweichende Position des Wortakzents zeigte sich in unserer Studie im Gegensatz zu den Befunden von Hirschfeld (1994) und Nossok (2009) als weniger störend.

Hypothesen zu den phonologischen Fehlerarten

Die in der **Hypothese 1.1** aufgestellte Annahme, dass suprasegmentale Abweichungen auf deutsche Muttersprachler eine störendere Wirkung haben als segmentale, konnte mit einem Signifikanzniveau von $p = 0,06$ nicht verifiziert werden. Dieses Ergebnis, das mit dem Befund von Hirschfeld (1994) kontrastiert, lässt sich nur durch die Zusammenstellung der Fehlerarten, die in der folgenden Tabelle präsentiert wird, begründen.

(82) Zusammenstellung der segmentalen und suprasegmentalen Fehlergruppen

Suprasegmentale Fehlerarten	Segmentale Fehlerarten
Gruppe C (Vokalquantität)	Gruppe A (Realisierung der Affrikate [ts])
Gruppe E (Konsonantenhäufung in der Wortanlautsilbe)	Gruppe B (Auslautverhärtung)
Gruppe F (Wortakzent)	Gruppe D (Realisierung des Umlauts)

Für die Überprüfung dieser Hypothese wurde der Gesamtmittelwert der suprasegmentalen mit dem der segmentalen Fehlerarten verglichen. Angesichts der Tatsache, dass die Gruppe D eindeutig schlechter als die anderen Gruppen bewertet wurde, können wir den unsignifikanten Unterschied auf den Einfluss der Bewertung dieser Gruppe auf den Gesamtmittelwert der segmentalen Gruppe zurückführen. Da der Umlaut nicht als rein phonologisches Element angesehen wird und der einzige Grund für die schlechte Bewertung der segmentalen Gruppe ist, bleibt fraglich, ob wir ohne ihn zu dem gleichen Ergebnis kommen würden. Außerdem ist anzumerken, dass dieses Resultat nur Aussagekraft für die Konstellation der in dieser Arbeit behandelten Fehlerarten haben kann. Hätten wir prosodische Erscheinungen wie Satzakzent oder Intonation als suprasegmentale Repräsentanten in unserer Studie einbezogen, wären wir vermutlich zu ganz anderen Ergebnissen gekommen.

In der **Hypothese 1.2** wurde angenommen, dass vokalische Abweichungen für deutsche Muttersprachler schwerwiegender auffallen als konsonantische. Für die Abweichungen im Vokalbereich berechneten wir den Mittelwert der Gruppen mit Fehlern in der Realisierung der Vokalquantität und des Umlauts, der hochsignifikant höher war als der Mittelwert der konsonantischen Gruppen mit der fehlerhaften Realisierung der Affrikate [ts] und der Verletzung der Auslautverhärtungsregel. Die vokalische Gruppe zeigte außerdem eine signifikant größere Differenz zu den korrekten Sätzen als die Gruppe der Konsonanten und somit ließ sich die Hypothese 1.2 definitiv bestätigen.

Die in der **Hypothese 1.3** vermutete Annahme, dass eine abweichende Realisierung der Vokallänge auf deutsche Muttersprachler störender wirkt als eine fehlerhafte Realisierung der Vokalqualität wurde für die Fehlerarten dieser Studie falsifiziert. Während die Abweichungen der Vokalquantität sich in den Untersuchungen von Hirschfeld (1994) und Nossok (2009) als besonders kommunikationsstörend erwiesen, wurde diese Fehlerart in unserer Arbeit signifikant besser bewertet als die abweichende Realisierung des Umlauts, die wir als Repräsentant der abweichenden Vokalqualität ausgewählt hatten. Diese Diskrepanz lässt sich, wie bei der Hypothese 1.1 beschrieben, nur auf die besonders störende Wirkung des abweichend ausgesprochenen Umlauts zurückführen.

Hypothesen zu den syntaktischen und morphosyntaktischen Fehlerarten

In der **Hypothese 2.1** wurde vermutet, dass eine fehlerhafte Genusmarkierung für deutsche Hörer schwerwiegender ist als eine falsche Realisierung des Kasus. Diese Hypothese konnte nicht bestätigt werden, weil kein signifikanter Unterschied zwischen diesen zwei Gruppen festgestellt wurde. Die annähernd gleich störende Wirkung beider Gruppen lässt sich vor allem dadurch begründen, dass die exakte Trennung bzw. Unterscheidung von Genus- und Kasusfehlern schwer realisierbar ist. Dies führt vor allem darauf zurück, dass das Kasusmerkmal hauptsächlich an den Artikeln zu erkennen ist.

Die in der **Hypothese 2.2** aufgestellte Annahme, dass ein Verbmorphologie-Fehler, der in unserer Arbeit als falsche Endung des Partizips II realisiert wurde, auf deutsche Muttersprachler störender wirkt als eine fehlerhafte Verbposition im Haupt- und Nebensatz, wurde nicht verifiziert. Das fehlerhafte Suffix des Partizips Perfekt wurde zwar mehr als alle anderen Fehlerarten als Abweichung erkannt und leicht schlechter als die falsche Verbposition im Hauptsatz bewertet, es konnten aber keine signifikanten Unterschiede zwischen der Bewertung dieser Fehlerart und der falschen Positionierung des Verbs festgestellt werden.

Die **Hypothese 2.3** ging davon aus, dass die falsche Verbposition im Hauptsatz für deutsche Hörer auffälliger bzw. störender klingt als die falsche Positionierung des Verbs im Nebensatz. Diese Annahme konnte angesichts der sogar leicht besseren Bewertung der fehlerhaften Verbposition im Hauptsatz nicht verifiziert werden. Da wir uns bei der Auswahl der Beispiele für die Position des Verbs im Hauptsatz nur auf die Verletzung einer Umstellungsregel, in der aufgrund des Vorkommens eines anderen Satzglieds in der ersten Position das Verb in der

zweiten und das Subjekt in der dritten Position steht, beschränkt haben, können wir die Ergebnisse dieser Studie nicht auf alle Ausprägungen der Verbposition im Hauptsatz übertragen. Auch unsere Aussagen über die Position des Verbs im Nebensatz beschränken sich nur auf die falsche Verbposition in Nebensätzen mit den eingeleiteten Konjunktionen *dass*, *wenn* und *obwohl*. Hätten wir z. B. Nebensätze mit *weil* als Konjunktion ausgewählt, hätten unsere Ergebnisse aufgrund der Existenz anderer umgangssprachlicher Abweichungen der Verbposition möglicherweise anders ausgesehen.

Hypothesen zum kontrastiven Vergleich der Grammatik- und Aussprachefehler

Anders als in der **Hypothese 3** erwartet, erwiesen sich die phonologischen Fehler signifikant weniger störend als die syntaktischen bzw. morphosyntaktischen. Die gegensätzlichen Ergebnisse dieser Arbeit und der anderen im Forschungsüberblick zitierten Untersuchungen lassen sich nur auf die Anzahl und die unterschiedlichen Ausspracheeigenheiten der Sprecher zurückführen. In allen erwähnten Studien im Kapitel 2.2.3., in denen sich die schlechte Aussprache als ein beeinträchtigender Faktor bei der Bewertung durch Muttersprachler erwies, wurde die Wirkung verschiedener Sprecher, die entweder unterschiedliche Akzentstufen hatten oder aus Muttersprachlern und Nichtmuttersprachlern bestanden, untersucht.

In unserer Arbeit beobachteten wir die unterschiedliche Wirkung der absichtlich gesprochenen Fehlerarten eines Sprechers, der Deutsch mit einem leichten Akzent sprach. Da unser Sprecher alle Sätze mit einem kontinuierlich leichten Akzent vorlas und seine Ausspracheeigenheit keinerlei Verständigungsprobleme verursachte, beschränken wir das Resultat dieser Hypothese nur auf die fortgeschrittenen DaF-Sprecher und behaupten, dass in den sprachlichen Äußerungen eines fortgeschrittenen DaF-Lerners, der deutlich, aber mit einem leichten ausländischen Akzent spricht, die grammatischen Abweichungen deutlich negativer auffallen als die phonologischen.

Basierend auf der Annahme, dass die Identifizierung und Benennung der Grammatikfehler für phonetisch ungeschulte Hörer einfacher wäre als die Erkennung der Aussprachefehler, sind wir in der **Hypothese 4.1** davon ausgegangen, dass grammatische Fehler eine höhere Anzahl der korrekten Erkennung aufweisen. Diese Hypothese wurde aufgrund des um 49% höheren Erkennungsgrads der grammatischen Abweichungen, was als hochsignifikanter Unterschied erwiesen wurde, definitiv verifiziert. Bei der Überprüfung dieser Hypothese beobachteten wir

zudem die Anzahl der falschen Erkennungen in beiden Gruppen und konnten dabei keine bedeutenden Unterschiede feststellen. Dafür zeigte sich die Anzahl der Probanden, die keine Angaben zu der Identifizierung der Fehler gemacht hatten, bei der Aussprache-Fehlergruppe deutlich höher als bei den Grammatikfehlerarten. Diese Tatsache zusammen mit dem Befund der dritten Hypothese, dass die phonologischen Fehlerarten eine positivere Bewertung aufweisen, können wir darauf zurückführen, dass phonologische Fehlerarten in unserer Arbeit deutlich weniger herausgehört wurden.

Die in der **Hypothese 4.2** aufgestellte Annahme, dass der Sicherheitsgrad bei der korrekten Erkennung der Grammatikfehler höher wäre als der der Aussprachefehler, wurde nicht eindeutig bestätigt. Die korrekt erkannten Grammatikfehler wurden von 82% der Probanden mit dem Sicherheitsgrad 1 bewertet, während diese Zahl bei den phonologischen Fehlerarten 68% betrug. Obwohl dieser Unterschied als signifikant erwiesen wurde, konnten wir die Hypothese nicht als definitiv bestätigt ansehen, da die berechnete Effektstärke sehr niedrig war. Wenn wir den großen Unterschied zwischen der Bewertung und der Anzahl der korrekten Erkennung des Fehlers beider Fehlergruppen betrachten und ihn mit dem Unterschied des Sicherheitsgrads vergleichen, können wir behaupten, dass der Sicherheitsgrad der korrekten Erkennung in beiden Fehlergruppen nicht bedeutend unterschiedlich war. Man kann hier davon ausgehen, dass diejenigen Probanden, die den Fehler tatsächlich erkannt haben, sich bei ihrer Erkennung sehr sicher waren, unabhängig davon, um welche Fehlerart es sich handelte.

Hypothesen zur allgemeinen Wahrnehmung der Sprechfehler

In der **Hypothese 5** war angenommen worden, dass zwischen der korrekten Erkennung des Fehlers und dessen negativer Bewertung ein positiver Zusammenhang besteht. Obwohl sich die berechnete Korrelation als signifikant erwies, konnten wir diese Hypothese nicht als eindeutig verifiziert betrachten, denn nur ca. 19% der schlechten Bewertungen konnten auf die korrekte Erkennung des Fehlers zurückgeführt werden. Um klarzustellen, welche anderen Faktoren für die weiteren 80% der schlechten Bewertung verantwortlich sind, beobachteten wir diesen Zusammenhang in verschiedenen Fehlergruppen.

Dabei stellte sich heraus, dass die Korrelation zwischen der Fehlererkennung und dessen negativer Bewertung in der Aussprache-Fehlergruppe eindeutig höher ist als in der Gruppe der Grammatikfehler. Darüber hinaus zeigte sich, dass die Korrelation bei den Fehlerarten, die

in der Bewertung besonders positiv abschnitten, höher ist als bei denjenigen, die schlechtere Bewertungen erhielten. Die äußerst negativ bewerteten Fehlergruppen, d. h. die Gruppen mit der falschen Realisierung des Partizips II und der Verbposition, wiesen sogar eine negative Korrelation auf. Man kann also die Ergebnisse der Analyse dieser Hypothese folgendermaßen zusammenfassen: Je besser die Fehlergruppe bewertet wurde, desto deutlicher zeigte sich der Zusammenhang zwischen der schlechten Bewertung und der korrekten Erkennung des Fehlers.

Dies bedeutet, dass bei positiv bewerteten Fehlergruppen der Mittelwertunterschied zwischen den erkannten und den nicht erkannten Fehlern größer war als bei den schlechter bewerteten Fehlerarten. Es ist also zu vermuten, dass bei den negativ evaluierten Fehlerarten die störende Wirkung des Fehlers so stark war, dass die Probanden auch die nicht erkannten Fehler unbewusst genauso schlecht und in manchen Fällen sogar schlechter bewertet haben als die erkannten Abweichungen. Die Hypothese 5 lässt sich insofern nur für die weniger störenden Fehlerarten bestätigen. Bei den negativer auffallenden Fehlerarten können außer der korrekten Erkennung andere Faktoren wie der schlechte Klang des fehlerhaften Satzes oder die unbewusste Empfindung des Fehlers für die schlechte Bewertung verantwortlich sein.

In der **Hypothese 6** sind wir davon ausgegangen, dass die korrekt gesprochenen Äußerungen eines Akzentsprechers fälschlicherweise als fehlerhaft bewertet werden. Der Gesamtmittelwert der korrekten Sätze von ($M = 1,49$) wies darauf hin, dass diese Sätze nicht hundertprozentig als standardsprachlich gut bewertet wurden. Für die Überprüfung dieser Hypothese schauten wir zusätzlich auf die Angaben der Probanden zu den erkannten Fehlern. Dabei stellte sich heraus, dass die korrekten Sätze zu einem großen Teil, nämlich ca. 88%, als fehlerfrei bewertet wurden. Die 12% der erkannten Fehler bezogen sich hauptsächlich auf vier Sätze, die jeweils von ca. 30% der Probanden als fehlerhaft gekennzeichnet wurden. Wenn wir unser Ergebnis mit den Resultaten anderer im Forschungsüberblick erwähnten Arbeiten wie der von Settinieri (2011:74) durchgeführten Untersuchung, in der ca. 28% der korrekten Sätze als fehlerhaft bewertet wurden, vergleichen und die Tatsache betrachten, dass unser Sprecher im Gegensatz zu ihren Sprechern ein Nichtmuttersprachler war, können wir behaupten, dass die fehlerfreien Sätze in unserer Arbeit tatsächlich als korrekt bewertet wurden. Diese Feststellung konnten wir jedoch nicht auf signifikante Unterschiede zurückführen, da wir keine vergleichbaren Ergebnisse haben, die die Bewertung der

Äußerungen eines Muttersprachlers oder eines Sprechers mit leichterem oder stärkerem Akzent reflektieren.

6. Diskussion und Schlussfolgerung

Das Hauptziel der vorliegenden Arbeit bestand darin, eine empirisch fundierte Basis für die Gewichtung und Klassifikation der gesprochenen Fehlerarten hinsichtlich ihres Störungsgrads für deutsche Rezipienten zu entwickeln. Um dieses Ziel zu erreichen, haben wir eine empirische Untersuchung durchgeführt, die die Beurteilung von 32 deutschen monolingualen Muttersprachlern über 12 verschiedene Fehlerarten reflektierte. Die Fehlerarten unterteilten sich in sechs phonologische und sechs grammatische Fehler, die eingebettet in Sätzen von einem Nichtmuttersprachler gesprochen wurden.

Betrachten wir unsere Untersuchungsmethode aus einer kritischen Perspektive, fällt an erster Stelle die Problematik auf, dass die Hörer aufgrund der Existenz eines fremden Akzents in der Aussprache des Sprechers vermutlich nicht nur die eingebauten Fehler, sondern auch andere Abweichungen negativ sanktioniert haben. Obwohl uns dieses Problem schon vor der Untersuchung bewusst war, haben wir einen Akzentsprecher ausgewählt, weil wir die Natürlichkeit der gesprochenen Sätze, die von einem DaF-Sprecher produziert werden sollten, beibehalten wollten. Um dieses Problem jedoch zu umgehen, entschieden wir uns für einen fortgeschrittenen DaF-Sprecher, der Deutsch mit einem leichten Akzent sprach. So gelang es uns, die Wahrscheinlichkeit, dass der Sprecher andere schwerwiegende Aussprachefehler beging, unter Kontrolle zu behalten.

Nicht zu vermeiden bei der auditiven Analyse war außerdem die subjektive Beurteilung der Probanden. Diese Subjektivität zeigt sich als ein unvermeidliches Problem bei allen Studien, die für die Datenerhebung Fragebögen oder Tests als Messinstrumente anwenden. Die Beantwortung der Frage, ob unsere Probanden beim Ankreuzen der Skalenpunkte tatsächlich die Fehlerhaftigkeit der Sätze oder die soziale Akzeptanz des Sprechers oder vielmehr seine Zugehörigkeit zu einer bestimmten Akzentgruppe bewertet haben, scheint nicht so einfach zu sein. Um dieser Frage nachzugehen und die Einstellung der Hörer gegenüber den zu untersuchenden Fehlerarten zu analysieren, befragten wir deshalb die Probanden in Form einer offenen Frage, welchen Fehler sie im Satz erkannt haben. Die einheitliche Bewertung der Sätze, die bei den gewonnenen deskriptiven Daten zu beobachten ist, kann als ein Beweis dafür dienen, dass wir bei den erzielten Ergebnissen tatsächlich von der Beurteilung der Fehlerhaftigkeit der gesprochenen Äußerungen ausgehen können.

Die Tatsache, dass einige Fehlerarten in unserer Studie ganz anders evaluiert wurden als in vergleichbaren Arbeiten, führten wir auf die Konstellation der Fehlerarten in unserer Untersuchung zurück. So schnitt z. B. im Gegensatz zu den Ergebnissen von Hirschfeld (1994) und Nossok (2009) in unserer Arbeit die falsche Positionierung des Wortakzents besonders positiv ab und es konnte kein signifikanter Unterschied zwischen den segmentalen und den suprasegmentalen Fehlerarten festgestellt werden. Aus diesem Grund beschränken wir unsere Ergebnisse nur auf die Zusammenstellung unserer Fehlerarten und gehen davon aus, dass eine weitere Untersuchung mit anderen Fehlerarten, z. B. im Bereich der Phonologie mit prosodischen Abweichungen in der Realisierung der Intonation oder des Satzakzents, zu unterschiedlichen Einschätzungen führen könnte.

Darüber hinaus haben unsere Ergebnisse nur Aussagekraft für die untersuchten Versuchspersonen, nämlich für junge Akademiker. Wären die Fehlerarten von einer anderen Hörergruppe, z. B. von einer Gruppe weniger gebildeter Hörer eines anderen Alters, bewertet worden, wären wir vermutlich zu ganz anderen Ergebnissen gekommen.

In Anbetracht der oben genannten Beschränkungen für die Aussagekräftigkeit unserer Ergebnisse wurde in der vorliegenden rezeptiven Untersuchung geschlussfolgert, dass die gesprochenen Fehlerarten, auch wenn sie keine Verständigungsprobleme verursachen, von deutschen Rezipienten unterschiedlich toleriert werden. Im Kontext unseres Experiments erwiesen sich die Grammatikfehler als eindeutig störender und auffälliger als die Aussprachefehler. Dieses Ergebnis können wir jedoch nur auf die geäußerten Fehler eines fortgeschrittenen DaF-Sprechers, der deutlich, aber mit einem leichten ausländischen Akzent spricht, beschränken. Es ist also anzunehmen, dass die Grammatikfehler vielleicht nicht so gravierend aufgefallen wären, hätte ein Sprecher mit stärkerem Akzent die fehlerhaften Sätze gesprochen. Diese Annahme kann vor allem auf die unterschiedliche Einstellung der Hörer gegenüber Sprechern mit verschiedenen Akzentstufen zurückgeführt werden, was in vielen Untersuchungen, aus denen in dieser Arbeit zitiert wurde, bestätigt werden konnte. Ein möglicher Grund für die besonders negative Bewertung der Grammatikfehler in unserer Arbeit wäre, dass die Hörer von einem fortgeschrittenen Sprecher erwartet hätten, keine solchen auffälligen Grammatikfehler zu begehen. Diese Interpretation bedarf allerdings einer weiteren parallelen Analyse, in der dieselben Fehler von einem Sprecher mit starkem Akzent produziert werden würden.

Die Resultate unserer Arbeit sollten vor allem DaF-Lernende und -Lehrkräfte für die Bedeutung der Sprechfehler für muttersprachliche Rezipienten sensibilisieren. Das Vorwissen darüber, welche Fehlerarten in der Wahrnehmung der Muttersprachler besonders sanktioniert auffallen, gibt den Lernenden und Lehrkräften die Möglichkeit, für die Vermeidung und Korrektur der Fehler Prioritäten zu setzen. Diese Gewichtung sollte natürlich nicht als Vernachlässigung der weniger störenden Fehler missverstanden werden. Aus eigener Erfahrung als DaF-Lernende und -Lehrkraft stelle ich allerdings fest, dass die Behebung der Fehler im DaF-Unterricht oft nicht im Einklang mit dem Störungsgrad ihrer Wirkung auf die muttersprachlichen Hörer steht. Ein Fehler wie die falsche Aussprache des Umlauts, die eine so gravierende Wirkung auf die deutschen Hörer ausübt, wird z. B. im DaF-Unterricht oft nicht genügend behandelt, während auf die Korrektur der Genusfehler, die laut unserer und der von Politzer (1978) erzielten Ergebnisse nicht so schwerwiegend auffallen, besonderen Wert gelegt wird.

Für die Konzeption und Entwicklung einer Richtlinie für die Fehlerkorrektur im DaF-Unterricht, an der sich sowohl die Lernenden als auch die Lehrkräfte orientieren können, ist an dieser Stelle auf folgende empirisch belegte Ergebnisse unserer Arbeit hinzuweisen:

- Die aussprachlichen Abweichungen, die einen leichten ausländischen Akzent reflektieren und die Verständigung nicht beeinträchtigen, werden von der Mehrheit der Deutschen besser toleriert als grammatische Inkorrektheiten.
- Die korrekte Aussprache des Umlauts, der Konsonantenhäufung in der Wortanlautsilbe und der Vokallänge ist für die Beurteilung der Aussprachekompetenz besonders wichtig, so dass deren Behandlung im DaF-Unterricht eine größere Bedeutung zugemessen werden muss.
- Die grammatischen Fehler bei der Positionierung des Verbs im Haupt- und Nebensatz sowie bei der Bildung des Partizips II fallen besonders störend auf, so dass auf deren Behandlung und Behebung im DaF-Unterricht mehr Wert gelegt werden muss.

Diese Schlussfolgerungen sind allerdings, wie vorhin erwähnt, nur für die fortgeschrittenen DaF-Sprecher generalisierbar. Um die Gültigkeit unserer Ergebnisse auszuweiten, sind zusätzliche empirische Untersuchungen vonnöten, in denen geprüft wird, wie unterschiedlich die Wirkung der sprachlichen Abweichungen bei anderen Akzentstufen aussieht. Darüber

hinaus wäre es interessant herauszufinden, wie eine andere Hörergruppe, z. B. eine Gruppe von Nicht-Akademikern einer anderen Altersklasse, diese Fehler beurteilen würde. Mögliche Fragestellungen für die künftige Forschung wären: Wie würden andere Fehlerarten oder sonstige Ausprägungen einer Fehlerart in einer ähnlichen Studie abschneiden? Würde die Untersuchung derselben Fehlerarten in Form einer anderen Sprachfertigkeit, z. B. als geschriebene Fehler, zu den gleichen Ergebnissen führen wie unsere? Wie würde die Gewichtung der Fehler aussehen, wenn man nicht die Fehlerhaftigkeit des Gesagten, sondern dessen soziale Akzeptanz untersucht hätte? Die Beantwortung dieser Fragen zusammen mit den Ergebnissen der vorliegenden Arbeit kann einen wesentlichen Beitrag dazu leisten, den Wunsch aller DaF-Lerner, fehlerfrei sprechen zu können, einigermaßen realisierbar werden zu lassen. Aus diesem Grund wäre es wünschenswert, wenn meine Arbeit als eine Basis für die weitere empirische Forschung über die Wirkung der sprachlichen Fehler fortgesetzt wird.

Literaturverzeichnis

- Adamcová, Livia (2002). Ausspracheschulung im Deutschunterricht. Analyse phonetischer Schwierigkeiten und Konsequenzen für die Unterrichtspraxis. *Sprache und Sprachen*, 29-30, 77-82.
- Albert, Ruth, & Koster, Cor J. (2002). *Empirie in Linguistik und Sprachlehrforschung. Ein methodologisches Arbeitsbuch*. Tübingen: Narr.
- Albert, Ruth, & Marx, Nicole (2014). *Empirisches Arbeiten in Linguistik und Sprachlehrforschung. Anleitung zu quantitativen Studien von der Planungsphase bis zum Forschungsbericht*. Tübingen: Narr (2. Aufl.).
- Andresen, H. (1976). Sprachnormen als Problem für Sprachunterricht und Sprachwissenschaft. *Osnabrücker Beiträge zur Sprachtheorie*, 1, 1-32.
- Anisfeld, Moshe, Bogo, Norman, & Lambert, Wallace E. (1962). Evaluational reactions to accented English speech. *Journal of Abnormal and Social Psychology*, 65, 223-231.
- Baird, Scott J. (1969). *Employment interview speech: A social dialect study in Austin, Texas*. University of Texas: Austin.
- Barkowski, Hans (2008). Integration und Sprache. Voraussetzungen und Grenzen der Unterstützung von Integrationsprozessen in Einwanderungsgesellschaften durch Maßnahmen zur Förderung des Zweitsprachenerwerbs der Immigrant/inn/en. In: Krumm, Hans-Jürgen, & Portmann-Tselikas, Paul R. (Hrsg.), *Theorie und Praxis. Österreichische Beiträge zu Deutsch als Fremdsprache*, 12/2008. Innsbruck, Wien & Bozen: Studienverlag.
- Barkowski, Hans, & Krumm, Hans-Jürgen (Hrsg.). (2010). *Fachlexikon Deutsch als Fremd- und Zweitsprache*. Tübingen & Basel: A. Francke.
- Bortz, Jürgen, & Schuster, Christof (2010). *Statistik für Human- und Sozialwissenschaftler*. Berlin & Heidelberg: Springer (7. Aufl.).
- Brennan, Eileen M., & Brennan, John S. (1981). Accent scaling and language attitudes: Reactions to Mexican American English speech. *Language and Speech*, 24/1, 209-221.
- Burt, Marina K. (1975). Error analysis in the adult EFL classroom. *TESOL Quarterly*, 9/1, 53-63.
- Bußmann, Hadumod (2008). *Lexikon der Sprachwissenschaft*. Stuttgart: Kröner.
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences*. Hillsdale, NJ: Erlbaum.

- Cohrs, Silke (2008). *Wirkung und Akzeptanz prosodischer Interferenzen Fremdsprachenlernender auf Muttersprachler. Dargestellt am Beispiel ungarischer Deutschlehrer*. Marburg: Tectum.
- Corder, S.P. (1978). The Significance of Learners Errors. In: Richards, J.C. (Hrsg.), *Error Analysis. Perspectives on second language acquisition*. (4. Aufl.). London: Longman.
- Coseriu, Eugenio (1975). *Sprachtheorie und allgemeine Sprachwissenschaft*. München: Wilhelm Fink.
- Delisle, Helga H. (1982). Native speaker judgment and the evaluation of errors in German. *The Modern Language Journal*, 66, 39-48.
- Duden, das Aussprachewörterbuch*. (2005). Bearb. von M. Mangold in Zusammenarbeit mit der Dudenredaktion. (6. Aufl., Bd. 6). Mannheim & Zürich: Dudenverlag.
- Field, Andy (2013). *Discovering statistics using IBM SPSS statistics: and sex and drugs and rock 'n' roll*. Los Angeles [u.a.]: SAGE.
- Hamidi, Zahra (2006). *Zur Phonologie des Persischen. System und Geschichte*. Dissertation. Universität München.
- Haug, Sonja (2008). *Sprachliche Integration von Migranten in Deutschland. Working Paper 14. Aus der Reihe „Integrationsreport“, Teil 2*. Nürnberg: Bundesamt für Migration und Flüchtlinge. Abgerufen von: <https://www.bamf.de/SharedDocs/Anlagen/DE/Publikationen/WorkingPapers/wp14-sprachliche-integration.pdf>. [12.03.2012]
- Heike, G. (1972). Die Rolle der Phonetik im Ausspracheunterricht. In: Nickel, Gerhard, & Raasch, Albert (Hrsg.), *Kongressbericht der 3. Jahrestagung der GAL*. Heidelberg: Julius Groos.
- Hellwig-Fábián, Inessa (2007). *Deutsch mit ausländischem Akzent. Eine empirische Studie zu Einstellungen junger Deutscher gegenüber Sprechern mit ostslavischer Muttersprache*. Frankfurt am Main: Peter Lang.
- Herbst, Thomas (1992). Pro-Nunciation. Zur Bedeutung einer guten Aussprache in der Fremdsprache. *Die Neueren Sprachen*, 91/1, 2-18.
- Hirschfeld, Ursula (1994). *Untersuchungen zur phonetischen Verständlichkeit Deutschlernender*. Frankfurt am Main: Hector.
- Hirschfeld, Ursula (1995). Phonetik im Unterricht Deutsch als Fremdsprache. Wie der Lehrer, so der Schüler?. *Fremdsprache Deutsch*, 12/1, 6-10.
- Hirschfeld, Ursula, Neuber, Baldur, & Stock, Eberhard (2010). Phonetische Sprechwirkungsforschung im Bereich der interkulturellen Kommunikation. In: Hirschfeld, U. et al. (Hrsg.), *Sprechwissenschaftlich-phonetische Untersuchungen*

- zur interkulturellen Kommunikation Russisch-Deutsch*. Frankfurt am Main [u.a.]: Peter Lang.
- Hohl, Kathrin (2007). *Ersetzungsmethoden für fehlende Werte kategorialer Variablen in klinischen Datensätzen*. Dissertation. Universität Ulm.
- Huber, J. (1974). Die traditionelle Sprachnorm und die Norm der kommunikativen Adäquanz. Zur Emanzipation von einer überflüssigen Norm in Wissenschaft und Schule. *Diskussion Deutsch*, 5, 144-153.
- Johansson, Stig (1973). The identification and evaluation of errors in foreign languages: A functional approach. In: Svartvik, Jan (Hrsg.), *Errata. Papers in error analysis*. Lund: CWK Gleerup.
- Katsikas, Sergios (1995). Fehler im Fremdspracherwerb: Vorschläge für eine Typologie von Fehlerursachen. *Moderne Sprachen*, 39/1-2, 19-34.
- Kaunzner, Ulrike A. (1997). *Aussprachekurs Deutsch. Ein komplettes Übungsprogramm zur Verbesserung der Aussprache für Unterricht und Selbststudium. Text- und Übungsbuch*. Heidelberg: Julius Groos.
- Kielhöfer, Bernd (1975). *Fehlerlinguistik des Fremdsprachenerwerbs. Linguistische, lernpsychologische und didaktische Analyse von Französischfehlern*. Kronberg/Ts: Scriptor.
- Kleppin, Karin (1998). *Fehler und Fehlerkorrektur. Fernstudieneinheit 19 des Fernstudienangebots „Germanistik/Deutsch als Fremdsprache“*. München: Langenscheidt.
- Kohler, Klaus J. (1995). *Einführung in die Phonetik des Deutschen*. Berlin: Erich Schmidt.
- Kolde, Gottfried (1980). Auswirkungen sprachlicher Fehler. In: Cherubim, D. (Hrsg.), *Fehlerlinguistik. Beiträge zum Problem der sprachlichen Abweichung*. Tübingen: Niemeyer, 172-188.
- Krech, Eva-Maria, Richter, Günther, Stock, Eberhard, & Suttner, Jutta (1991). *Sprechwirkung. Grundfragen, Methoden und Ergebnisse ihrer Erforschung*. Berlin: Akademie Verlag.
- Krech, Eva-Maria (Hrsg.). (1964). *Wörterbuch der deutschen Aussprache*. Leipzig: Enzyklopädie.
- Kuckartz, Udo, Rädiker, Stefan, Ebert, Thomas, & Schehl, Julia (2013). *Statistik. Eine verständliche Erklärung*. Wiesbaden: VS Springer.
- Kürschner, Wilfried (2008). *Grammatisches Kompendium. Systematisches Verzeichnis grammatischer Grundbegriffe*. Tübingen & Basel: Francke.

- Lambert, W.E., Hodgson, R.C., Gardner, R.C., & Fillenbaum, S. (1960). Evaluational reactions to spoken languages. *Journal of Abnormal and Social Psychology*, 60/1, 44-51.
- Lasswell, Harold D., Lerner, Daniel, & Pool, Ithiel de Sola (1952). *The comparative study of symbols*. Stanford: Stanford University Press.
- Lochner, Susanne, Büttner, Tobias, & Schuller, Karin (2013). *Das Integrationspanel. Langfristige Integrationsverläufe von ehemaligen Teilnehmenden an Integrationskursen. Working Paper 52*. Nürnberg: Bundesamt für Migration und Flüchtlinge. Abgerufen von: <http://www.bamf.de/SharedDocs/Anlagen/DE/Publikationen/WorkingPapers/wp52-abschluss-integrationspanel.pdf>. [12.03.2012]
- Messingschlager, Martin (2012). *Fehlende Werte in den Sozialwissenschaften. Analyse und Korrektur mit Beispielen aus dem ALLBUS*. Bamberg: University of Bamberg Press.
- Nossok, Swetlana (2009). *Kontrastive phonologische und phonetische Analyse Weißrussisch-Deutsch und Analyse interferenzbedingter Ausspracheabweichungen*. Frankfurt am Main: Peter Lang.
- Pazan, Atoosa (2008). *Wortakzent im Persischen und im Deutschen*. Unveröffentlichte Bachelor-Arbeit. Universität Marburg.
- Pazan, Atoosa (2011). *Prosodische Fehler bei persischsprachigen Deutschlernern*. Unveröffentlichte Master-Arbeit. Universität Marburg.
- Politzer, Robert L. (1978). Errors of English speakers of German as perceived and evaluated by German natives. *The Modern Language Journal*, 62, 253-261.
- Porst, Rolf (2014). *Fragebogen. Ein Arbeitsbuch*. Wiesbaden: Springer (4. Aufl.).
- Raabe, Horst (1980). Der Fehler beim Fremdsprachenerwerb und Fremdsprachengebrauch. In: Cherubim, D. (Hrsg.), *Fehlerlinguistik. Beiträge zum Problem der sprachlichen Abweichung*. Tübingen: Niemeyer, 61-93.
- Raab-Steiner, Elisabeth, & Benesch, Michael (2008). *Der Fragebogen. Von der Forschungsidee zur SPSS-Auswertung*. Wien: Facultas
- Raisler, I. (1976). Differential response to the same message delivered by native and foreign speakers. *Foreign Language Annals*, 9, 256-259.
- Ramers, Karl-Heinz (1998). *Einführung in die Phonologie*. München: Fink.
- Rausch, R.I. (1988). *Deutsche Phonetik für Ausländer. Ein Lehr- und Übungsbuch*. Leipzig: VEB Verlag Enzyklopädie.

- Reinke, Kerstin (2008). Zur Wirkung phonetischer Mittel in sachlich intendierter Sprechweise bei Deutsch sprechenden Russen. In: Anders, L.C. et al. (Hrsg.), *Hallesche Schriften zur Sprechwissenschaft und Phonetik*, 26. Frankfurt am Main: Peter Lang.
- Ryan, Ellen B., Carranza, Miguel A., & Moffie, Robert W. (1977). Reactions toward varying degrees of accentedness in the speech of Spanish-English bilinguals. *Language and Speech*, 20/3, 267-273.
- Schäfer, Thomas (2010). *Statistik I. Deskriptive und explorative Datenanalyse*. Wiesbaden: VS Springer.
- Schäfer, Thomas (2011). *Statistik II. Inferenzstatistik*. Wiesbaden: VS Springer.
- Schuller, Karin, Lochner, Susanne, Rother, Nina, & Hörner, Denise (2012). *Das Integrationspanel. Entwicklung der Deutschkenntnisse und Fortschritte der Integration bei Teilnehmenden an Alphabetisierungskursen. Working Paper 42*. Nürnberg: Bundesamt für Migration und Flüchtlinge. Abgerufen von: <https://www.bamf.de/SharedDocs/Anlagen/DE/Publikationen/WorkingPapers/wp42-integrationspanel.pdf>. [12.03.2012]
- Settinieri, Julia (2011). Soziale Akzeptanz unterschiedlicher Normabweichungen in der L2-Aussprache Deutsch. *Zeitschrift für interkulturellen Fremdsprachenunterricht. Didaktik und Methodik im Bereich Deutsch als Fremdsprache*, 16/2, 66-80.
- Spillner, Bernd (1991). *Error analysis. A comprehensive bibliography*. Amsterdam & Philadelphia: John Benjamins Publishing Company.
- SPSS Base 16.0 Benutzerhandbuch*. (2007). Chicago, IL. Abgerufen von: http://www.uni-muenster.de/imperia/md/content/ziv/service/software/spss/handbuecher/deutsch/spss_base_benutzerhandbuch_16.0.pdf. [06.01.2015]
- Szarska-Wieruszewska, Joanna (2005). *Lernerstrategien als Grundlage der lernertypischen Fehlerbehandlungen. Forschungsstudien am Beispiel der Deutsch lernenden, polnischen Studenten*. Hamburg: Verlag Dr. Kovač.
- Tucker, G. Richard, & Sarofim, Marian (1979). Investigating linguistic acceptability with Egyptian EFL students. *TESOL Quarterly*, 13/1, 29-39.
- Wiese, Richard (1987). Phonologie und Morphologie des Umlauts im Deutschen. *Zeitschrift für Sprachwissenschaft*, 6/2, 227-248.
- Wiese, Richard (1996). *The Phonology of German*. Oxford: Clarendon Press.
- Wiese, Richard (2011). *Phonetik und Phonologie*. Paderborn: Wilhelm Fink.

Anhang

Anhang 1	Testmaterialien	A.1
Anhang 2	Angaben zu den einzelnen Hörern	A.6
Anhang 3	Fragebögen zur Beurteilung der gesprochenen Sätze	A.7
Anhang 4	Codeplan für variablen	A.8
Anhang 5	Ergebnisse der Kolmogorov-Smirnov-Tests	A.9
	5.A.1. Ergebnisse der K-S-Tests für die fehlerhaften Sätze	A.9
	5.A.2. Ergebnisse der K-S-Tests für die korrekten Sätze	A.12
	5.B. Ergebnisse der K-S-Tests für die Fehlergruppen und die parallelen Gruppen mit korrekten Sätzen	A.13
	5.C. Ergebnisse der K-S-Tests für die Differenz-Variablen	A.15
	5.D. Ergebnisse der K-S-Tests für die segmentalen und suprasegmentalen Fehlergruppen	A.15
	5.E. Ergebnisse der K-S-Tests für die vokalischen und konsonantischen Fehlergruppen	A.16
	5.F. Ergebnisse des K-S-Tests für die Gruppe (K+L)	A.16
	5.G. Ergebnisse der K-S-Tests für die Aussprache- und Grammatik-Fehlergruppen	A.17
	5.H. Ergebnisse der K-S-Tests für die Sicherheitsangaben der Aussprache- und Grammatik-Fehlergruppen	A.17
Anhang 6	Deskriptive Ergebnisse für die einzelnen korrekten Sätze.....	A.18
Anhang 7	Sicherheitsgradtabellen für die einzelnen Fehlergruppen	A.20
Anhang 8	Vergleichende deskriptive Daten für die Aussprache- und Grammatik-Fehlerarten.....	A.24
Anhang 9	Paarweiser Vergleich der Differenzgruppen.....	A.26
Anhang 10	Testergebnisse für die Hypothese 2.2.....	A.28
	10.A. Wilcoxon- und T-Test-Ergebnisse.....	A.28

	10.B. Friedman-Test-Ergebnisse.....	A.29
Anhang 11	Kontrastiver Vergleich der fehlerhaften Satzpaare anhand der deskriptiven Werte und des Wilcoxon-Tests.....	A.31
Anhang 12	Kreuztabellen für den McNemar-Test.....	A.33
	12.A. Kreuztabelle für die Fehlerangaben der Grammatik- und Aussprache-Fehlergruppen.....	A.33
	12.B. Kreuztabelle für die Sicherheitsangaben der Grammatik- und Aussprache-Fehlergruppen.....	A.34
Anhang 13	Fehlerangaben in den einzelnen korrekten Sätzen.....	A.35

Anhang 1: Testmaterialien

A. Realisierung der Affrikate [ts]

1. Ich habe leider keine Zeit.
1. A. Ich habe leider keine [s]eit.
1. G. Ich habe leider *keinen* Zeit. (Genus)

2. Das Justizlexikon steht im Regal.
2. A. Das *Justi[s]lexikon* steht im Regal.
2. H. Das Justizlexikon steht *ins* Regal. (Kasus)

3. Er denkt nicht an seine Zukunft.
3. A. Er denkt nicht an seine [s]ukunft.
3. I. Er denkt nicht *über* seine Zukunft. (Präposition))

4. Meine Augen waren vom Glanz der Sonne geblendet.
4. A. Meine Augen waren vom *Glan[s]* der Sonne geblendet.
4. J. Meine Augen waren vom Glanz der Sonne *geblenden*. (Partizip2)

5. Neulich ist Benzin viel teurer geworden.
5. A. Neulich ist *Ben[s]in* viel teurer geworden.
5. K. Neulich *Benzin ist* viel teurer geworden. (Verb im HS)

6. Alle wissen, dass die Touristen sich mit Malaria infiziert haben.
6. A. Alle wissen, dass die Touristen sich mit Malaria *infi[s]iert* haben.
6. L. Alle wissen, dass die Touristen *haben* sich mit Malaria infiziert. (Verb im NS)

B. Auslautverhärtung

7. Meine Freundin ist naiv, aber nicht dumm.
7. B. Meine Freundin ist *nai[v]*, aber nicht dumm.
7. G. *Mein* Freundin ist naiv, aber nicht dumm. (Genus)

8. Der Wind blies mir ins Gesicht.

8. B. Der Wind *blie[z]* mir ins Gesicht.

8. H. Der Wind blies mir *im* Gesicht. (Kasus)

9. Er beteiligt sich immer aktiv am Unterricht.

9. B. Er beteiligt sich immer *akti[v]* am Unterricht.

9. I. Er beteiligt sich immer aktiv *zum* Unterricht. (Präposition)

10. Einer der Hunde war sehr aggressiv und hat andere Hunde verletzt.

10. B. Einer der Hunde war sehr *agressi[v]* und hat andere Hunde verletzt.

10. J. Einer der Hunde war sehr aggressiv und hat andere Hunde *verletzen*. (Partizip 2)

11. Leider wies die Rechnung einige Fehler auf.

11. B. Leider *wie[z]* die Rechnung einige Fehler auf.

11. K. Leider *die Rechnung wies* einige Fehler auf. (Verb im HS)

12. Wenn du dich mehr informieren willst, lies dieses Buch.

12. B. Wenn du dich mehr informieren willst, *lie[z]* dieses Buch.

12. L. Wenn du *willst* dich mehr informieren, lies dieses Buch. (Verb im NS)

C. Vokalquantität

13. Wie lange kannst du deinen Atem anhalten?

13. C. Wie lange kannst du deinen *[a]tem* anhalten?

13. G. Wie lange kannst du *deine* Atem anhalten? (Genus)

14. Sie ist die gute Seele in unserem Team.

14. C. Sie ist die gute *S[e]le* in unserem Team.

14. H. Sie ist die gute Seele in *unser* Team. (Kasus)

15. Mein Zahn ist sehr empfindlich und tut beim Kauen weh.

15. C. Mein *Z[a]n* ist sehr empfindlich und tut beim Kauen weh.

15. I. Mein Zahn ist sehr empfindlich und tut *am* Kauen weh. (Präposition)
16. Es wird ewig dauern, bis er seine Schulden bezahlt hat.
16. C. Es wird [*e*]wig dauern, bis er seine Schulden bezahlt hat.
16. J. Es wird ewig dauern, bis er seine Schulden *bezahlen* hat. (Partizip 2)
17. Am Wochenende sollte das Wetter schön werden.
17. C. Am Wochenende sollte das W[*ε*:]tter schön werden.
17. K. Am Wochenende *das Wetter sollte* schön werden. (Verb im HS)
18. Die Kinder spielen mit dem Ball, obwohl sie es gar nicht dürfen.
18. C. Die Kinder spielen mit dem B[*a*:]ll, obwohl sie es gar nicht dürfen.
18. L. Die Kinder spielen mit dem Ball, obwohl sie *dürfen* es gar nicht. (Verb im NS)

D. Realisierung des Umlauts

19. Die Gastgeber servierten ein üppiges Mahl.
19. D. Die Gastgeber servierten ein [*ʊ*]ppiges Mahl.
19. G. Die Gastgeber servierten *eine üppige* Mahl. (Genus)
20. Kein Problem. Das nehme ich dir nicht übel.
20. D. Kein Problem. Das nehme ich dir nicht [*u*:]bel.
20. H. Kein Problem. Das nehme ich *dich* nicht übel. (Kasus)
21. Sie sind sehr gütig zu mir.
21. D. Sie sind sehr g[*u*:]tig zu mir.
21. I. Sie sind sehr gütig *mit* mir. (Präposition)
22. Sie hat sich gar nicht bemüht.
22. D. Sie hat sich gar nicht bem[*u*:]ht.
22. J. Sie hat sich gar nicht *bemühen*. (Partizip2)

23. Jeden Tag treffen sich die Künstler.

23. D. Jeden Tag treffen sich die *K[ʊ]nstler*.

23. K. Jeden Tag *die Künstler treffen sich*. (Verb im HS)

24. Auch wenn es hier sehr gemütlich ist, müssen wir bald wieder aufbrechen.

24. D. Auch wenn es hier sehr *gem[u:]tlich* ist, müssen wir bald wieder aufbrechen.

24. L. Auch wenn es hier *ist* sehr gemütlich, müssen wir bald wieder aufbrechen. (Verb im NS)

E. Konsonantenhäufung in der Wortanlautsilbe

25. Deutsch ist für mich eine zweite Sprache.

25. E. Deutsch ist für mich eine zweite *[e]Sprache*.

25. G. Deutsch ist für mich *ein zweites* Sprache. (Genus)

26. Wir sind auf der Straße.

26. E. Wir sind auf der *[e]Straße*.

26. H. Wir sind auf *die* Straße. (Kasus)

27. Ich habe Angst vor Schlangen.

27. E. Ich habe Angst vor *[e]Schlangen*.

27. I. Ich habe Angst *gegen* Schlangen. (Präposition)

28. Die Verbrennungen haben ihr unerträgliche Schmerzen bereitet.

28. E. Die Verbrennungen haben ihr unerträgliche *[e]Schmerzen* bereitet.

28. J. Die Verbrennungen haben ihr unerträgliche Schmerzen *bereiten*. (Partizip2)

29. Letztes Jahr habe ich mein Studium beendet.

29. E. Letztes Jahr habe ich mein *[e]Studium* beendet.

29. K. Letztes Jahr *ich habe* mein Studium beendet. (Verb im HS)

30. Obwohl er nicht so groß ist, kann er drei Meter hoch springen.

30. E. Obwohl er nicht so groß ist, kann er drei Meter hoch *[e]springen*.

30. L. Obwohl er *ist* nicht so groß, kann er drei Meter hoch springen. (Verb im NS)

F. Wortakzent

31. Wir treffen uns am Hauptbahnhof.

31. F. Wir treffen uns am *Haupt'bahnhof*.

31. G. Wir treffen uns *an der* Hauptbahnhof. (Genus)

32. Die Kinder sind auf dem Spielplatz.

32. F. Die Kinder sind auf dem *Spiel'platz*.

32. H. Die Kinder sind auf *den* Spielplatz. (Kasus)

33. Meine Freundin ist immer noch im Krankenhaus.

33. F. Meine Freundin ist immer noch im *Kranken'haus*.

33. I. Meine Freundin ist immer noch *auf* dem Krankenhaus. (Präposition)

34. Der Taxifahrer hat an der Bushaltestelle gehalten.

34. F. Der Taxifahrer hat an der *Bus'haltestelle* gehalten.

34. J. Der Taxifahrer hat an der Bushaltestelle *gehalten*. (Partizip2)

35. Als ausländischer Student muss man in Deutschland zuerst einen Sprachkurs besuchen.

35. F. Als ausländischer Student muss man in Deutschland zuerst einen Sprach'*kurs* besuchen.

35. K. Als ausländischer Student *man muss* in Deutschland zuerst einen Sprachkurs besuchen. (Verb im HS)

36. Ich wusste nicht, dass das Studienkolleg hier in der Nähe ist.

36. F. Ich wusste nicht, dass das Studien'*kolleg* hier in der Nähe ist.

36. L. Ich wusste nicht, dass das Studienkolleg *ist* hier in der Nähe. (Verb im NS)

Anhang 2: Angaben zu den einzelnen Hörern

Versuchspersonen	Geschlecht	Alter	Studium/Ausbildung	Beruf
1	W	U30	Europäische Ethnologie/Kulturwissenschaften und Soziologie (FB 3)	Bildungsreferentin in einer entwicklungspolitischen NGO
2	M	U30	Medienwissenschaft (FB 9)	
3	M	Ü30	Sinologie (FB 10)	
4	W	Ü30	Lehrerin	Lehrerin
5	M	U30	MA Religionswissenschaft (FB 3)	
6	M	U30	BA Bildungs- und Erziehungswissenschaften (FB 21)	
7	M	U30	Lehramt Sport/Geschichte (FB 6&21)	
8	M	U30	BA Mathematik (FB 12)	
9	W	Ü30	Diplom Angewandte Fremdsprachen und Wirtschaft (FB 2&10)	Lehrkraft (Englisch)
10	W	U30	Studium (FB 09&10)	
11	W	U30	Lehramt (FB 10&6)	
12	M	Ü30		Dozent und Jobcoach in der Erwachsenenbildung
13	M	U30	Humanmedizin (FB 20)	
14	W	U30	Studium der Orientwissenschaft mit dem Schwerpunkt Politik (FB 3 &10)	
15	M	U30	Humanmedizin (FB 20)	
16	M	Ü30		Dozent und Jobcoach in der Erwachsenenbildung
17	M	U30	Philosophie (FB 3)	
18	W	U30	Politikwissenschaft (FB 3)	
19	M	U30	Geschichte, Sport (FB 6 & 21)	
20	M	U30	BA Geschichte (FB 6)	
21	M	U30	BA Sozialwissenschaften (FB 3)	
22	M	U30	Lehramt Französisch/Spanisch (FB 10)	
23	M	U30	Lehramt Germanistik/Geschichte (FB 6&9)	
24	W	U30	Sozialwissenschaften (FB 3)	
25	W	U30	Lehramt Englisch/Ethik (FB 10&3)	
26	M	U30	Lehramt Gymnasium Politik/ Wirtschaft/Ethik/Philosophie/ Deutsch (FB 3&9)	
27	M	Ü30	Lehramt (FB 21)	Pädagoge
28	M	U30	Politikwissenschaft (FB 3)	
29	W	U30	Lehramt Spanisch/Ethik/DaF (FB 10, 9&3)	
30	W	U30	Pharmazie (FB 16)	Apothekerin/ wissenschaftliche Mitarbeiterin
31	M	U30	MA Geschichte der internationalen Politik (FB 6)	
32	W	U30	Sozialwissenschaften (FB 3)	Bürokauffrau + Sozialwissenschaftlerin

Anhang 3: Fragebögen zur Beurteilung der gesprochenen Sätze

Einige Angaben zu Ihrer Person

Geschlecht: weiblich ☐ männlich ☐

Alter: 18 bis 30 ☐ über 30 ☐

Ausbildung oder Studium: _____

Beruf: _____

Ihre Beurteilung

Hören Sie sich bitte die gesprochenen Sätze an. Versuchen Sie beim Zuhören die Deutschkenntnisse des Sprechers einzuschätzen, indem Sie für jeden Satz ankreuzen, wie gut oder wie schlecht der Satz auf Standard-/Hochdeutsch gesprochen wurde. Dafür haben Sie eine sechsstufige Skala.

Falls Sie ein bestimmtes Wort oder einen bestimmten Ausdruck als Fehler erkennen, können Sie es aufschreiben und angeben, wie sicher Sie sich bei dieser Erkennung sind.

Wie gut wurde der Satz auf Standarddeutsch gesprochen?

Satz 1.

sehr gut 1 2 3 4 5 6 sehr schlecht

Fehler: _____ sicher 1 2 3 4 unsicher

.

.

.

Satz 108.

sehr gut 1 2 3 4 5 6 sehr schlecht

Fehler: _____ sicher 1 2 3 4 unsicher

Anhang 4: Codeplan für Variablen

Variablen-name	Variablenlabel	Wertelabels	Messniveau	fehlende Werte
S (1-108)	Beurteilung der Sätze (1-108)	1=sehr gut 2=2 2,5=2,5 3=3 3,5=3,5 4=4 5=5 6=sehr schlecht	intervall	99= wenn korrekter Satz schlechter beurteilt wurde als die falsche Version
F (1-108)	Fehlerangaben für Sätze (1-108)	0=keine Angabe 1=Fehler richtig erkannt 2=fälschlicherweise Aussprachefehler gehört 3=fälschlicherweise Grammatikfehler gehört 4=fälschlicherweise Stil-/Wortschatz- oder andere Fehler gehört 5=als unverständlich beurteilt	nominal	9= irrelevante Angaben wie Rechtschreibfehler
Si (1-108)	Sicherheitsgrad für die Erkennung des Fehlers in Sätzen(1-108)	1=sicher 1,5=1,5 2=2 2,5=2,5 3=3 4=unsicher	intervall	999= keine Angabe

Anmerkungen

1. Folgende Bewertungen wurden unter anderem als richtige Erkennung des Fehlers registriert:

A. Die Anmerkungen, bei denen das fehlerhafte Wort aufgeschrieben wurde, auch wenn es keine Hinweise gab, um was für einen Fehler es sich handelte.

B. Wenn bei einem Satz, der Grammatikfehler beinhaltete, nur das Wort „Grammatik“ und für die Aussprachefehler nur das Wort „Aussprache“ eingetragen wurde.

C. Wenn jemand bei den Fehlern der Vokalquantität die falsche Betonung des Wortes als Fehler angegeben hat.

2. Im Satz S.17: *soll* statt *sollte* wurde als Stilfehler dokumentiert.

Anhang 5: Ergebnisse der Kolmogorov-Smirnov-Tests

Anhang 5.A.1.: Ergebnisse der K-S-Tests für die fehlerhaften Sätze

		Satz1.A	Satz2.A	Satz3.A	Satz4.A	Satz5.A	Satz6.A
H		31	28	30	26	28	31
Parameter der	Mittelwert	2,097	1,393	1,867	1,538	2,143	2,000
Normalverteilung	Standardabweichung	,8309	,6853	,8604	,8115	1,0079	,8563
Extremste Differenzen	absolut	,288	,395	,243	,362	,199	,210
	positiv	,288	,395	,243	,362	,199	,210
	negativ	-,228	-,283	-,162	-,253	-,160	-,177
Teststatistik		,288	,395	,243	,362	,199	,210
Asymp. Sig. (2-seitig)		,000	,000	,000	,000	,006	,001

		Satz7.B	Satz8.B	Satz9.B	Satz10.B	Satz11.B	Satz12.B
H		29	30	14	27	27	27
Parameter der	Mittelwert	1,276	1,367	1,429	2,000	1,778	1,407
Normalverteilung	Standardabweichung	,5276	,6149	,6462	1,1435	,8916	,6939
Extremste Differenzen	absolut	,458	,424	,389	,291	,290	,388
	positiv	,458	,424	,389	,291	,290	,388
	negativ	-,301	-,276	-,254	-,191	-,191	-,279
Teststatistik		,458	,424	,389	,291	,290	,388
Asymp. Sig. (2-seitig)		,000	,000	,000	,000	,000	,000

		Satz13.C	Satz14.C	Satz15.C	Satz16.C	Satz17.C	Satz18.C
H		29	24	24	29	29	32
Parameter der	Mittelwert	2,000	1,500	2,542	2,241	2,414	1,625
Normalverteilung	Standardabweichung	1,0351	,5898	1,0206	,9124	1,0183	,7071
Extremste Differenzen	absolut	,293	,343	,327	,211	,244	,312
	positiv	,293	,343	,327	,191	,244	,312
	negativ	-,167	-,260	-,214	-,211	-,170	-,202
Teststatistik		,293	,343	,327	,211	,244	,312
Asymp. Sig. (2-seitig)		,000	,000	,000	,002	,000	,000

		Satz19.D	Satz20.D	Satz21.D	Satz22.D	Satz23.D	Satz24.D
H		32	32	32	32	31	31
Parameter der	Mittelwert	2,531	2,469	2,750	2,594	2,613	2,419
Normalverteilung	Standardabweichung	1,0468	,9499	1,0776	,9108	1,0223	,8860
Extremste Differenzen	absolut	,225	,314	,226	,274	,306	,359
	positiv	,225	,314	,226	,274	,306	,359
	negativ	-,150	-,217	-,181	-,195	-,210	-,253
Teststatistik		,225	,314	,226	,274	,306	,359
Asymp. Sig. (2-seitig)		,000	,000	,000	,000	,000	,000

		Satz25.E	Satz26.E	Satz27.E	Satz28.E	Satz29.E	Satz30.E
H		28	31	31	30	31	31
Parameter der	Mittelwert	2,964	3,000	2,161	1,400	2,226	2,645
Normalverteilung	Standardabweichung	1,1701	1,0954	,7788	,7240	1,2030	1,0181
Extremste Differenzen	absolut	,238	,206	,246	,410	,252	,221
	positiv	,238	,206	,195	,410	,252	,221
	negativ	-,155	-,148	-,246	-,290	-,154	-,166
Teststatistik		,238	,206	,246	,410	,252	,221
Asymp. Sig. (2-seitig)		,000	,002	,000	,000	,000	,001

		Satz31.F	Satz32.F	Satz33.F	Satz34.F	Satz35.F	Satz36.F
H		25	29	32	31	28	31
Parameter der	Mittelwert	1,480	1,759	1,813	1,806	1,929	1,806
Normalverteilung	Standardabweichung	,7141	,8305	1,0607	,8725	,8133	,8334
Extremste Differenzen	absolut	,389	,302	,278	,274	,230	,253
	positiv	,389	,302	,278	,274	,230	,253
	negativ	-,251	-,180	-,222	-,178	-,192	-,172
Teststatistik		,389	,302	,278	,274	,230	,253
Asymp. Sig. (2-seitig)		,000	,000	,000	,000	,001	,000

		Satz1.G	Satz7.G	Satz13.G	Satz19.G	Satz25.G	Satz31.G
H		32	32	29	32	30	31
Parameter der	Mittelwert	2,594	2,063	2,483	3,375	3,533	2,806
Normalverteilung	Standardabweichung	1,1601	,9136	,9495	1,0999	1,1059	1,0139
Extremste Differenzen	absolut	,196	,215	,246	,196	,252	,206
	positiv	,196	,215	,246	,196	,252	,206
	negativ	-,137	-,160	-,168	-,153	-,181	-,156
Teststatistik		,196	,215	,246	,196	,252	,206
Asymp. Sig. (2-seitig)		,003	,001	,000	,003	,000	,002

		Satz2.H	Satz8.H	Satz14.H	Satz20.H	Satz26.H	Satz32.H
H		32	32	32	32	32	29
Parameter der	Mittelwert	3,344	2,906	2,656	3,031	3,156	1,759
Normalverteilung	Standardabweichung	1,1807	1,0883	,8273	1,0313	1,0809	,9876
Extremste Differenzen	absolut	,211	,266	,317	,262	,276	,331
	positiv	,195	,266	,317	,262	,276	,331
	negativ	-,211	-,203	-,214	-,159	-,161	-,221
Teststatistik		,211	,266	,317	,262	,276	,331
Asymp. Sig. (2-seitig)		,001	,000	,000	,000	,000	,000

		Satz3.I	Satz9.I	Satz15.I	Satz21.I	Satz27.I	Satz33.I
H		32	31	29	28	32	32
Parameter der	Mittelwert	3,031	2,871	3,345	1,786	3,188	2,891
Normalverteilung	Standardabweichung	,9995	1,0565	1,1425	1,0313	1,0298	1,1621
Extremste Differenzen	absolut	,231	,279	,205	,313	,228	,213
	positiv	,231	,279	,205	,313	,228	,213
	negativ	-,151	-,205	-,131	-,223	-,147	-,162
Teststatistik		,231	,279	,205	,313	,228	,213
Asymp. Sig. (2-seitig)		,000	,000	,003	,000	,000	,001

		Satz4.J	Satz10.J	Satz16.J	Satz22.J	Satz28.J	Satz34.J
H		31	30	32	30	32	32
Parameter der	Mittelwert	2,935	3,400	2,813	3,367	2,813	3,500
Normalverteilung	Standardabweichung	1,1236	1,0372	1,0298	,9994	,9311	1,2181
Extremste Differenzen	absolut	,251	,283	,303	,276	,277	,191
	positiv	,251	,283	,303	,276	,277	,172
	negativ	-,203	-,183	-,184	-,190	-,191	-,191
Teststatistik		,251	,283	,303	,276	,277	,191
Asymp. Sig. (2-seitig)		,000	,000	,000	,000	,000	,005

		Satz5.K	Satz11.K	Satz17.K	Satz23.K	Satz29.K	Satz35.K
H		31	30	32	31	32	31
Parameter der	Mittelwert	3,484	3,233	3,031	3,290	3,000	2,839
Normalverteilung	Standardabweichung	1,2348	1,2507	1,1496	1,0706	1,0473	,9344
Extremste Differenzen	absolut	,176	,238	,284	,188	,236	,267
	positiv	,176	,238	,284	,188	,236	,267
	negativ	-,148	-,162	-,185	-,166	-,170	-,185
Teststatistik		,176	,238	,284	,188	,236	,267
Asymp. Sig. (2-seitig)		,016	,000	,000	,007	,000	,000

		Satz6.L	Satz12.L	Satz18.L	Satz24.L	Satz30.L	Satz36.L
H		32	32	31	31	31	32
Parameter der	Mittelwert	3,531	3,438	3,226	3,000	3,323	3,281
Normalverteilung	Standardabweichung	1,1067	1,1053	1,1750	,9309	1,0766	1,0846
Extremste Differenzen	absolut	,247	,216	,189	,213	,187	,196
	positiv	,247	,216	,189	,213	,181	,196
	negativ	-,159	-,132	-,133	-,149	-,187	-,152
Teststatistik		,247	,216	,189	,213	,187	,196
Asymp. Sig. (2-seitig)		,000	,001	,006	,001	,007	,003

Anhang 5.A.2.: Ergebnisse der K-S-Tests für die korrekten Sätze

		Satz1	Satz2	Satz3	Satz4	Satz5	Satz6
H		32	32	32	32	32	32
Parameter der	Mittelwert	1,156	1,344	1,313	1,531	1,781	1,469
Normalverteilung	Standardabweichung	,3689	,6016	,5351	,7177	,8701	,6713
Extremste Differenzen	absolut	,508	,435	,439	,364	,284	,383
	positiv	,508	,435	,439	,364	,284	,383
	negativ	-,336	-,284	-,280	-,230	-,185	-,242
Teststatistik		,508	,435	,439	,364	,284	,383
Asymp. Sig. (2-seitig)		,000	,000	,000	,000	,000	,000

		Satz7	Satz8	Satz9	Satz10	Satz11	Satz12
H		32	32	32	32	32	32
Parameter der	Mittelwert	1,281	1,344	2,031	1,625	1,625	1,438
Normalverteilung	Standardabweichung	,5811	,7007	,8224	,8707	,8328	,6690
Extremste Differenzen	absolut	,467	,438	,234	,357	,336	,400
	positiv	,467	,438	,234	,357	,336	,400
	negativ	-,314	-,312	-,204	-,236	-,226	-,257
Teststatistik		,467	,438	,234	,357	,336	,400
Asymp. Sig. (2-seitig)		,000	,000	,000	,000	,000	,000

		Satz13	Satz14	Satz15	Satz16	Satz17	Satz18
H		32	32	32	32	32	32
Parameter der	Mittelwert	1,625	1,438	2,313	1,656	1,375	1,281
Normalverteilung	Standardabweichung	,7071	,5644	1,2032	,7453	,7071	,5811
Extremste Differenzen	absolut	,312	,375	,321	,311	,421	,467
	positiv	,312	,375	,321	,311	,421	,467
	negativ	-,202	-,247	-,179	-,189	-,298	-,314
Teststatistik		,312	,375	,321	,311	,421	,467
Asymp. Sig. (2-seitig)		,000	,000	,000	,000	,000	,000

		Satz19	Satz20	Satz21	Satz22	Satz23	Satz24
H		32	32	32	32	32	32
Parameter der	Mittelwert	1,438	1,313	1,375	1,250	1,688	1,594
Normalverteilung	Standardabweichung	,6189	,5923	,6091	,5080	,7803	,8370
Extremste Differenzen	absolut	,385	,451	,418	,470	,311	,355
	positiv	,385	,451	,418	,470	,311	,355
	negativ	-,240	-,299	-,269	-,311	-,189	-,239
Teststatistik		,385	,451	,418	,470	,311	,355
Asymp. Sig. (2-seitig)		,000	,000	,000	,000	,000	,000

		Satz25	Satz26	Satz27	Satz28	Satz29	Satz30
H		32	32	32	32	32	32
Parameter der	Mittelwert	2,188	1,438	1,406	1,188	1,313	1,703
Normalverteilung	Standardabweichung	1,1483	,6690	,6148	,3966	,5923	,7055
Extremste Differenzen	absolut	,224	,400	,402	,494	,451	,278
	positiv	,224	,400	,402	,494	,451	,278
	negativ	-,167	-,257	-,254	-,318	-,299	-,226
Teststatistik		,224	,400	,402	,494	,451	,278
Asymp. Sig. (2-seitig)		,000	,000	,000	,000	,000	,000

		Satz31	Satz32	Satz33	Satz34	Satz35	Satz36
H		32	32	32	32	32	32
Parameter der	Mittelwert	1,469	1,375	1,188	1,281	1,500	1,375
Normalverteilung	Standardabweichung	,6713	,8707	,4709	,6342	,6720	,5536
Extremste Differenzen	absolut	,383	,448	,498	,484	,365	,407
	positiv	,383	,448	,498	,484	,365	,407
	negativ	-,242	-,333	-,345	-,329	-,228	-,249
Teststatistik		,383	,448	,498	,484	,365	,407
Asymp. Sig. (2-seitig)		,000	,000	,000	,000	,000	,000

Anhang 5.B.: Ergebnisse der K-S-Tests für die Fehlergruppen und die parallelen Gruppen mit korrekten Sätzen

		Mittelwert der Gr. A	Mittelwert der korrekten Sätze Gr.A	Mittelwert der Gr. B	Mittelwert der korrekten Sätze Gr.B	Mittelwert der Gr. C	Mittelwert der korrekten Sätze Gr.C
H		32	32	32	32	32	32
Parameter der	Mittelwert	1,8646	1,4323	1,6057	1,5573	2,0635	1,6146
Normalverteilung	Standardabweichung	,65122	,43350	,64410	,55618	,67243	,54450
Extremste	absolut	,182	,188	,178	,188	,112	,177
Differenzen	positiv	,182	,188	,178	,188	,112	,177
	negativ	-,121	-,159	-,174	-,158	-,106	-,130
Teststatistik		,182	,188	,178	,188	,112	,177
Asymp. Sig. (2-seitig)		,009	,006	,011	,006	,200	,012

		Mittelwert der Gr. D	Mittelwert der korrekten Sätze Gr.D	Mittelwert der Gr. E	Mittelwert der korrekten Sätze Gr.E	Mittelwert der Gr. F	Mittelwert der korrekten Sätze Gr.F
H		32	32	32	32	32	32
Parameter der	Mittelwert	2,5583	1,4427	2,3833	1,5391	1,7818	1,3646
Normalverteilung	Standardabweichung	,84578	,46875	,80731	,46518	,69668	,50701
Extremste	absolut	,184	,186	,151	,127	,178	,246
Differenzen	positiv	,184	,186	,151	,127	,178	,246
	negativ	-,102	-,172	-,074	-,123	-,131	-,236
Teststatistik		,184	,186	,151	,127	,178	,246
Asymp. Sig. (2-seitig)		,008	,006	,060	,200	,012	,000

		Mittelwert der Gr. G	Mittelwert der korrekten Sätze Gr.G	Mittelwert der Gr. H	Mittelwert der korrekten Sätze Gr.H	Mittelwert der Gr. I	Mittelwert der korrekten Sätze Gr.I
H		32	32	32	32	32	32
Parameter der	Mittelwert	2,7875	1,5260	2,8385	1,3750	2,8651	1,6042
Normalverteilung	Standardabweichung	,80245	,46588	,87269	,47330	,90880	,51631
Extremste	absolut	,124	,131	,125	,214	,160	,174
Differenzen	positiv	,124	,131	,109	,201	,160	,174
	negativ	-,081	-,129	-,125	-,214	-,097	-,121
Teststatistik		,124	,131	,125	,214	,160	,174
Asymp. Sig. (2-seitig)		,200	,173	,200	,001	,037	,015

		Mittelwert der Gr. J	Mittelwert der korrekten Sätze Gr.J	Mittelwert der Gr. K	Mittelwert der korrekten Sätze Gr.K	Mittelwert der Gr. L	Mittelwert der korrekten Sätze Gr.L
H		32	32	32	32	32	32
Parameter der	Mittelwert	3,1240	1,4219	3,1328	1,5469	3,2958	1,4766
Normalverteilung	Standardabweichung	,93939	,52872	,99333	,52404	,93375	,46722
Extremste	absolut	,121	,279	,165	,223	,156	,190
Differenzen	positiv	,121	,279	,165	,223	,156	,190
	negativ	-,085	-,212	-,127	-,148	-,061	-,154
Teststatistik		,121	,279	,165	,223	,156	,190
Asymp. Sig. (2-seitig)		,200	,000	,027	,000	,046	,005

Anhang 5.C.: Ergebnisse der K-S-Tests für die Differenz-Variablen

		Differenz A-KA	Differenz B-KB	Differenz C-KC	Differenz D-KD	Differenz E-KE	Differenz F-KF
H		32	32	32	32	32	32
Parameter der	Mittelwert	,4323	,0484	,4490	1,1156	,8443	,4172
Normalverteilung	Standardabweichung	,47076	,24678	,40919	,63911	,58241	,50996
Extremste	absolut	,122	,098	,139	,187	,195	,188
Differenzen	positiv	,122	,090	,080	,187	,195	,188
	negativ	-,085	-,098	-,139	-,116	-,111	-,113
Teststatistik		,122	,098	,139	,187	,195	,188
Asymp. Sig. (2-seitig)		,200	,200	,121	,006	,003	,005

		Differenz G-KG	Differenz H-KH	Differenz I-KI	Differenz J-KJ	Differenz K-KK	Differenz L-KL
H		32	32	32	32	32	32
Parameter der	Mittelwert	1,2615	1,4635	1,2609	1,7021	1,5859	1,8193
Normalverteilung	Standardabweichung	,56073	,80234	,76590	,89747	,81374	,72442
Extremste	absolut	,098	,169	,159	,129	,167	,146
Differenzen	positiv	,098	,169	,159	,129	,167	,146
	negativ	-,069	-,094	-,113	-,090	-,107	-,105
Teststatistik		,098	,169	,159	,129	,167	,146
Asymp. Sig. (2-seitig)		,200	,020	,039	,188	,023	,081

Anhang 5.D.: Ergebnisse der K-S-Tests für die segmentalen und suprasegmentalen Fehlergruppen

		Mittelwert der segmentalen Fehlergruppen	Mittelwert der suprasegmentalen Fehlergruppen	Mittelwert der Differenzen der segmentalen Fehlergruppen	Mittelwert der Differenzen der suprasegmentalen Fehlergruppen
H		32	32	32	32
Parameter der	Mittelwert	2,0095	2,0762	,5321	,5701
Normalverteilung	Standardabweichung	,66298	,67171	,34995	,40800
Extremste	absolut	,195	,157	,211	,134
Differenzen	positiv	,195	,157	,211	,134
	negativ	-,123	-,089	-,117	-,130
Teststatistik		,195	,157	,211	,134
Asymp. Sig. (2-seitig)		,003	,044	,001	,150

Anhang 5.E.: Ergebnisse der K-S-Tests für die vokalischen und konsonantischen Fehlergruppen

		Mittelwert der Bewertungen der Vokal- Fehlergruppen (C+D)	Mittelwert der Bewertungen der Konsonanten- Fehlergruppen (A+B)	Mittelwert der Differenz- Variablen C und D	Mittelwert der Differenz- Variablen A und B
H		32	32	32	32
Parameter der	Mittelwert	2,3109	1,7352	,7823	,2404
Normalverteilung	Standardabweichung	,70006	,62578	,41636	,27373
Extremste Differenzen	absolut	,165	,218	,141	,128
	positiv	,165	,218	,141	,128
	negativ	-,084	-,149	-,070	-,118
Teststatistik		,165	,218	,141	,128
Asymp. Sig. (2-seitig)		,026	,001	,106	,197

Anhang 5.F.: Ergebnisse des K-S-Tests für die Gruppe (K+L)

		Mittelwert K und L	Mittelwert der Differenzen der Gruppen K und L
H		32	32
Parameter der Normalverteilung	Mittelwert	3,2143	1,7026
	Standardabweichung	,94871	,74736
Extremste Differenzen	absolut	,156	,165
	positiv	,156	,165
	negativ	-,100	-,100
Teststatistik		,156	,165
Asymp. Sig. (2-seitig)		,047	,027

Anhang 5.G.: Ergebnisse der K-S-Tests für die Aussprache- und Grammatik-Fehlergruppen

		Mittelwert der Phonologie- Fehlergruppe	Mittelwert der Grammatik- Fehlergruppe	Mittelwert der Differenzen der Phonologie- Fehlergruppe	Mittelwert der Differenzen der Grammatik- Fehlergruppe
H		32	32	32	32
Parameter der	Mittelwert	2,0429	3,0073	,5511	1,5155
Normalverteilung	Standardabweichung	,65623	,85624	,34724	,67650
Extremste Differenzen	absolut	,145	,133	,119	,139
	positiv	,145	,133	,119	,139
	negativ	-,103	-,101	-,103	-,067
Teststatistik		,145	,133	,119	,139
Asymp. Sig. (2-seitig)		,083	,159	,200	,120

Anhang 5.H.: Ergebnisse des K-S-Tests für die Sicherheitsangaben der Aussprache- und Grammatik-Fehlergruppe

		Sicherheitsangaben bei der Erkennung der Aussprachefehler	Sicherheitsangaben bei der Erkennung der Grammatikfehler
H		648	902
Parameter der Normalverteilung	Mittelwert	1,458	1,234
	Standardabweichung	,7612	,5511
Extremste Differenzen	absolut	,405	,483
	positiv	,405	,483
	negativ	-,274	-,335
Teststatistik		,405	,483
Asymp. Sig. (2-seitig)		,000	,000

Anhang 6: Deskriptive Ergebnisse für die einzelnen korrekten Sätze

	Satz1	Satz2	Satz3	Satz4	Satz5	Satz6
N gültig	32	32	32	32	32	32
fehlend	0	0	0	0	0	0
Mittelwert	1,156	1,344	1,313	1,531	1,781	1,469
Median	1,000	1,000	1,000	1,000	2,000	1,000
Modalwert	1,0	1,0	1,0	1,0	1,0	1,0
Standardabweichung	,3689	,6016	,5351	,7177	,8701	,6713

	Satz7	Satz8	Satz9	Satz10	Satz11	Satz12
N gültig	32	32	32	32	32	32
fehlend	0	0	0	0	0	0
Mittelwert	1,281	1,344	2,031	1,625	1,625	1,438
Median	1,000	1,000	2,000	1,000	1,000	1,000
Modalwert	1,0	1,0	2,0	1,0	1,0	1,0
Standardabweichung	,5811	,7007	,8224	,8707	,8328	,6690

	Satz13	Satz14	Satz15	Satz16	Satz17	Satz18
N gültig	32	32	32	32	32	32
fehlend	0	0	0	0	0	0
Mittelwert	1,625	1,438	2,313	1,656	1,375	1,281
Median	1,500	1,000	2,000	1,500	1,000	1,000
Modalwert	1,0	1,0	2,0	1,0	1,0	1,0
Standardabweichung	,7071	,5644	1,2032	,7453	,7071	,5811

	Satz19	Satz20	Satz21	Satz22	Satz23	Satz24
N gültig	32	32	32	32	32	32
fehlend	0	0	0	0	0	0
Mittelwert	1,438	1,313	1,375	1,250	1,688	1,594
Median	1,000	1,000	1,000	1,000	1,500	1,000
Modalwert	1,0	1,0	1,0	1,0	1,0	1,0
Standardabweichung	,6189	,5923	,6091	,5080	,7803	,8370

	Satz25	Satz26	Satz27	Satz28	Satz29	Satz30
N gültig	32	32	32	32	32	32
fehlend	0	0	0	0	0	0
Mittelwert	2,188	1,438	1,406	1,188	1,313	1,703
Median	2,000	1,000	1,000	1,000	1,000	2,000
Modalwert	1,0	1,0	1,0	1,0	1,0	1,0
Standardabweichung	1,1483	,6690	,6148	,3966	,5923	,7055

	Satz31	Satz32	Satz33	Satz34	Satz35	Satz36
N gültig	32	32	32	32	32	32
fehlend	0	0	0	0	0	0
Mittelwert	1,469	1,375	1,188	1,281	1,500	1,375
Median	1,000	1,000	1,000	1,000	1,000	1,000
Modalwert	1,0	1,0	1,0	1,0	1,0	1,0
Standardabweichung	,6713	,8707	,4709	,6342	,6720	,5536

Anhang 7: Sicherheitsgradtabellen für die einzelnen Fehlergruppen

Sicherheitsangaben Gruppe A

		Häufigkeit	Prozent	gültige Prozent	kumulative Prozent
gültig	sicher	24	45,3	61,5	61,5
	2	9	17,0	23,1	84,6
	3	6	11,3	15,4	100,0
	Gesamtsumme	39	73,6	100,0	
fehlend	999,0	14	26,4		
Gesamtsumme		53	100,0		

Sicherheitsangaben Gruppe B

		Häufigkeit	Prozent	gültige Prozent	kumulative Prozent
gültig	2	2	66,7	100,0	100,0
fehlend	999,0	1	33,3		
Gesamtsumme		3	100,0		

Sicherheitsangaben Gruppe C

		Häufigkeit	Prozent	gültige Prozent	kumulative Prozent
gültig	sicher	19	48,7	59,4	59,4
	2	10	25,6	31,3	90,6
	3	3	7,7	9,4	100,0
	Gesamtsumme	32	82,1	100,0	
fehlend	999,0	7	17,9		
Gesamtsumme		39	100,0		

Sicherheitsangaben Gruppe D

		Häufigkeit	Prozent	gültige Prozent	kumulative Prozent
gültig	sicher	121	73,8	87,1	87,1
	2	13	7,9	9,4	96,4
	3	3	1,8	2,2	98,6
	unsicher	2	1,2	1,4	100,0
	Gesamtsumme	139	84,8	100,0	
fehlend	999,0	25	15,2		
Gesamtsumme		164	100,0		

Sicherheitsangaben Gruppe E

		Häufigkeit	Prozent	gültige Prozent	kumulative Prozent
gültig	sicher	32	41,0	50,8	50,8
	2	21	26,9	33,3	84,1
	3	8	10,3	12,7	96,8
	unsicher	2	2,6	3,2	100,0
	Gesamtsumme	63	80,8	100,0	
fehlend	999,0	15	19,2		
Gesamtsumme		78	100,0		

Sicherheitsangaben Gruppe F

		Häufigkeit	Prozent	gültige Prozent	kumulative Prozent
gültig	sicher	16	32,7	42,1	42,1
	2	17	34,7	44,7	86,8
	3	3	6,1	7,9	94,7
	unsicher	2	4,1	5,3	100,0
	Gesamtsumme	38	77,6	100,0	
fehlend	999,0	11	22,4		
Gesamtsumme		49	100,0		

Sicherheitsangaben Gruppe G

		Häufigkeit	Prozent	gültige Prozent	kumulative Prozent
gültig	sicher	87	68,0	83,7	83,7
	2	13	10,2	12,5	96,2
	3	3	2,3	2,9	99,0
	unsicher	1	,8	1,0	100,0
	Gesamtsumme	104	81,3	100,0	
fehlend	999,0	24	18,8		
Gesamtsumme		128	100,0		

Sicherheitsangaben Gruppe H

		Häufigkeit	Prozent	gültige Prozent	kumulative Prozent
gültig	sicher	115	72,3	84,6	84,6
	2	16	10,1	11,8	96,3
	3	4	2,5	2,9	99,3
	unsicher	1	,6	,7	100,0
	Gesamtsumme	136	85,5	100,0	
fehlend	999,0	23	14,5		
Gesamtsumme		159	100,0		

Sicherheitsangaben Gruppe I

		Häufigkeit	Prozent	gültige Prozent	kumulative Prozent
gültig	sicher	111	73,0	88,8	88,8
	2	11	7,2	8,8	97,6
	3	2	1,3	1,6	99,2
	unsicher	1	,7	,8	100,0
	Gesamtsumme	125	82,2	100,0	
fehlend	999,0	27	17,8		
Gesamtsumme		152	100,0		

Sicherheitsangaben Gruppe J

		Häufigkeit	Prozent	gültige Prozent	kumulative Prozent
gültig	sicher	129	74,1	86,6	86,6
	2	16	9,2	10,7	97,3
	3	4	2,3	2,7	100,0
	Gesamtsumme	149	85,6	100,0	
fehlend	999,0	25	14,4		
Gesamtsumme		174	100,0		

Sicherheitsangaben Gruppe K

		Häufigkeit	Prozent	gültige Prozent	kumulative Prozent
gültig	sicher	117	71,3	86,7	86,7
	2	15	9,1	11,1	97,8
	3	3	1,8	2,2	100,0
	Gesamtsumme	135	82,3	100,0	
fehlend	999,0	29	17,7		
Gesamtsumme		164	100,0		

Sicherheitsangaben Gruppe L

		Häufigkeit	Prozent	gültige Prozent	kumulative Prozent
gültig	sicher	112	67,9	82,4	82,4
	1,5	1	,6	,7	83,1
	2	17	10,3	12,5	95,6
	3	5	3,0	3,7	99,3
	unsicher	1	,6	,7	100,0
	Gesamtsumme	136	82,4	100,0	
fehlend	999,0	29	17,6		
Gesamtsumme		165	100,0		

Anhang 8: Vergleichende deskriptive Daten für die Aussprache- und Grammatik-Fehlerarten

Deskriptive Werte der phonologischen Fehlergruppen

	Mittelwert der Gruppe A	Mittelwert der Gruppe B	Mittelwert der Gruppe C	Mittelwert der Gruppe D	Mittelwert der Gruppe E	Mittelwert der Gruppe F
N gültig	32	32	32	32	32	32
fehlend	0	0	0	0	0	0
Mittelwert	1,8646	1,6057	2,0635	2,5583	2,3833	1,7818
Standardfehler des Mittelwerts	,11512	,11386	,11887	,14951	,14271	,12316
Median	1,6333	1,4000	2,0000	2,4167	2,3333	1,6333
Modalwert	1,17	1,00	1,50	2,00	2,00	1,00
Standardabweichung	,65122	,64410	,67243	,84578	,80731	,69668

Deskriptive Werte der grammatischen Fehlergruppen

	Mittelwert der Gruppe G	Mittelwert der Gruppe H	Mittelwert der Gruppe I	Mittelwert der Gruppe J	Mittelwert der Gruppe K	Mittelwert der Gruppe L
N gültig	32	32	32	32	32	32
fehlend	0	0	0	0	0	0
Mittelwert	2,7875	2,8385	2,8651	3,1240	3,1328	3,2958
Standardfehler des Mittelwerts	,14185	,15427	,16065	,16606	,17560	,16506
Median	2,7500	2,6667	2,6667	2,9167	2,8750	3,2500
Modalwert	2,00	1,83	2,17	2,00	2,17	2,67
Standardabweichung	,80245	,87269	,90880	,93939	,99333	,93375

Deskriptive Werte der Differenz-Variablen der phonologischen Fehlergruppen

	Differenz A- KA	Differenz B- KB	Differenz C- KC	Differenz D- KD	Differenz E- KE	Differenz F- KF
N gültig	32	32	32	32	32	32
fehlend	0	0	0	0	0	0
Mittelwert	,4323	,0484	,4490	1,1156	,8443	,4172
Standardfehler des Mittelwerts	,08322	,04363	,07234	,11298	,10296	,09015
Median	,3333	,0167	,5000	1,0000	,7500	,2500
Modalwert	,17	-,17	,33	,83	,33	,00
Standardabweichung	,47076	,24678	,40919	,63911	,58241	,50996

Deskriptive Werte der Differenz-Variablen der grammatischen Fehlergruppen

		Differenz G- KG	Differenz H- KH	Differenz I- KI	Differenz J- KJ	Differenz K- KK	Differenz L- KL
N	gültig	32	32	32	32	32	32
	fehlend	0	0	0	0	0	0
Mittelwert		1,2615	1,4635	1,2609	1,7021	1,5859	1,8193
Standardfehler des Mittelwerts		,09912	,14184	,13539	,15865	,14385	,12806
Median		1,1667	1,2500	1,0833	1,5000	1,3333	1,6667
Modalwert		1,83	1,00	,83	1,00	1,00	1,17
Standardabweichung		,56073	,80234	,76590	,89747	,81374	,72442

Anhang 9: Paarweiser Vergleich der Differenzgruppen

Paarweiser Vergleich der Differenzgruppen der phonologischen Fehlergruppen

Stichprobe1-Stichprobe2	Test-statistik	Std. Fehler	Standard Test-statistik	Sig.	Angep. Sig.
Differenz B-KB-Differenz F-KF	-1,281	,468	-2,739	,006	,092
Differenz B-KB-Differenz A-KA	1,328	,468	2,840	,005	,068
Differenz B-KB-Differenz C-KC	-1,609	,468	-3,441	,001	,009
Differenz B-KB-Differenz E-KE	-3,156	,468	-6,748	,000	,000
Differenz B-KB-Differenz D-KD	-3,688	,468	-7,884	,000	,000
Differenz F-KF-Differenz A-KA	,047	,468	,100	,920	1,000
Differenz F-KF-Differenz C-KC	,328	,468	,702	,483	1,000
Differenz F-KF-Differenz E-KE	1,875	,468	4,009	,000	,001
Differenz F-KF-Differenz D-KD	2,406	,468	5,145	,000	,000
Differenz A-KA-Differenz C-KC	-,281	,468	-,601	,548	1,000
Differenz A-KA-Differenz E-KE	-1,828	,468	-3,909	,000	,001
Differenz A-KA-Differenz D-KD	-2,359	,468	-5,045	,000	,000
Differenz C-KC-Differenz E-KE	-1,547	,468	-3,307	,001	,014
Differenz C-KC-Differenz D-KD	-2,078	,468	-4,443	,000	,000
Differenz E-KE-Differenz D-KD	,531	,468	1,136	,256	1,000

Paarweiser Vergleich der Differenzgruppen der grammatischen Fehlergruppen

Stichprobe1-Stichprobe2	Test-statistik	Std. Fehler	Standard Test-statistik	Sig.	Angep. Sig.
Differenz I-KI-Differenz G-KG	,062	,468	,134	,894	1,000
Differenz I-KI-Differenz H-KH	,719	,468	1,537	,124	1,000
Differenz I-KI-Differenz K-KK	-1,172	,468	-2,506	,012	,183
Differenz I-KI-Differenz J-KJ	-1,484	,468	-3,174	,002	,023
Differenz I-KI-Differenz L-KL	-2,094	,468	-4,477	,000	,000
Differenz G-KG-Differenz H-KH	-,656	,468	-1,403	,161	1,000
Differenz G-KG-Differenz K-KK	-1,109	,468	-2,372	,018	,265
Differenz G-KG-Differenz J-KJ	-1,422	,468	-3,040	,002	,035
Differenz G-KG-Differenz L-KL	-2,031	,468	-4,343	,000	,000
Differenz H-KH-Differenz K-KK	-,453	,468	-,969	,333	1,000
Differenz H-KH-Differenz J-KJ	-,766	,468	-1,637	,102	1,000
Differenz H-KH-Differenz L-KL	-1,375	,468	-2,940	,003	,049
Differenz K-KK-Differenz J-KJ	,312	,468	,668	,504	1,000
Differenz K-KK-Differenz L-KL	-,922	,468	-1,971	,049	,731
Differenz J-KJ-Differenz L-KL	-,609	,468	-1,303	,193	1,000

Anhang 10: Testergebnisse für die Hypothese 2.2.

Anhang 10.A.: Wilcoxon- und T-Test-Ergebnisse

Wilcoxon-Test-Ergebnis für K & J

	Mittelwert der Gruppe K - Mittelwert der Gruppe J
U	-,091 ^b
Asymp. Sig. (2-seitig)	,927

Wilcoxon-Test-Ergebnis für L & J

	Mittelwert der Gruppe L - Mittelwert der Gruppe J
U	-1,651 ^b
Asymp. Sig. (2-seitig)	,099

Wilcoxon-Test-Ergebnis für (K+L) & J

	Mittelwert K und L - Mittelwert der Gruppe J
U	-1,051
Asymp. Sig. (2-seitig)	,293

Wilcoxon-Test-Ergebnis für die Differenz- Variablen K & J

	Differenz K-KK - Differenz J-KJ
U	-1,013
Asymp. Sig. (2-seitig)	,311

Wilcoxon-Test-Ergebnis für die Differenz- Variablen (K+L) & J

	Mittelwert der Differenzen der Gruppen K und L - Differenz J-KJ
U	-,432
Asymp. Sig. (2-seitig)	,666

T-Test-Ergebnis für die Differenz-Variablen L & J

	Paarige Differenzen					t	df	Sig. (2-seitig)
	Mittelwert	Standardabweichung	Standardfehler Mittelwert	95% Konfidenzintervall der				
				Differenz				
				Unterer	Oberer			
Paar 1 Differenz J-KJ - Differenz L-KL	-,11719	,60582	,10709	-,33561	,10123	-1,094	31	,282

Anhang 10.B.: Friedman-Test-Ergebnisse

Friedman-Test-Ergebnis für die Gruppen K, L, J und (K+L)

Ränge

	Mittlerer Rang
Mittelwert der Gruppe J	2,36
Mittelwert der Gruppe K	2,13
Mittelwert der Gruppe L	2,94
Mittelwert K und L	2,58

Teststatistiken

H	32
Chi-Quadrat	7,404
df	3
Asymp. Sig.	,060

Friedman-Test-Ergebnis für die Differenz-Variablen der Gruppen K, L, J und (K+L)

Ränge

	Mittlerer Rang
Differenz J-KJ	2,31
Differenz K-KK	1,97
Differenz L-KL	3,13
Mittelwert der Differenzen der Gruppen K und L	2,59

Teststatistiken

H	32
Chi-Quadrat	14,299
df	3
Asymp. Sig.	,003

Post-hoc-Ergebnis für die Differenz-Variablen der Gruppen K, L, J und (K+L)

Stichprobe1-Stichprobe2	Test-statistik	Std. Fehler	Standard Test-statistik	Sig.	Angep. Sig.
Differenz J-KJ-Mittelwert der Differenzen der Gruppen K und L	-,281	,323	-,871	,384	1,000
Differenz K-KK-Differenz J-KJ	,344	,323	1,065	,287	1,000
Mittelwert der Differenzen der Gruppen K und L-Differenz L-KL	,531	,323	1,646	,100	,599
Differenz K-KK-Mittelwert der Differenzen der Gruppen K und L	-,625	,323	-1,936	,053	,317
Differenz J-KJ-Differenz L-KL	-,812	,323	-2,517	,012	,071
Differenz K-KK-Differenz L-KL	-1,156	,323	-3,583	,000	,002

Anhang 11: Kontrastiver Vergleich der fehlerhaften Satzpaare anhand der deskriptiven Werte und des Wilcoxon-Tests

	Satz1.A	Satz1.G	Satz2.A	Satz2.H	Satz3.A	Satz3.I	Satz4.A	Satz4.J	Satz5.A	Satz5.K	Satz6.A	Satz6.L
N gültig	31	32	28	32	30	32	26	31	28	31	31	32
fehlend	1	0	4	0	2	0	6	1	4	1	1	0
Mittelwert	2,097	2,594	1,393	3,344	1,867	3,031	1,538	2,935	2,143	3,484	2,000	3,531
Median	2,000	2,500	1,000	3,500	2,000	3,000	1,000	3,000	2,000	3,000	2,000	3,000
Standard-abweichung	,8309	1,1601	,6853	1,1807	,8604	,9995	,8115	1,1236	1,0079	1,2348	,8563	1,1067

	Satz1.G – Satz1.A ¹	Satz2.H – Satz2.A	Satz3.I – Satz3.A	Satz4.J – Satz4.A	Satz5.K – Satz5.A	Satz6.L – Satz6.A
Z	-2,878b	-4,323b	-3,882b	-3,988b	-4,323b	-4,633b
Asymp.Sig. (2-seitig)	,004	,000	,000	,000	,000	,000

	Satz7. B	Satz7. G	Satz8. B	Satz8. H	Satz9. B	Satz9. I	Satz10. B	Satz10. J	Satz11. B	Satz11. K	Satz12. B	Satz12. L
N gültig	29	32	30	32	14	31	27	30	27	30	27	32
fehlend	3	0	2	0	18	1	5	2	5	2	5	0
Mittelwert	1,276	2,063	1,367	2,906	1,429	2,871	2,000	3,400	1,778	3,233	1,407	3,438
Median	1,000	2,000	1,000	3,000	1,000	3,000	2,000	3,000	2,000	3,000	1,000	3,000
Standard-abweichung	,5276	,9136	,6149	1,0883	,6462	1,0565	1,1435	1,0372	,8916	1,2507	,6939	1,1053

	Satz7.G – Satz7.B	Satz8.H – Satz8.B	Satz9.I – Satz9.B	Satz10.J – Satz10.B	Satz11.K – Satz11.B	Satz12.L – Satz12.B
Z	-3,601b	-4,747b	-3,022b	-3,690b	-4,218b	-4,435b
Asymp. Sig. (2-seitig)	,000	,000	,003	,000	,000	,000

	Satz13. C	Satz13. G	Satz14. C	Satz14. H	Satz15. C	Satz15. I	Satz16. C	Satz16. J	Satz17. C	Satz17. K	Satz18. C	Satz18. L
N gültig	29	29	24	32	24	29	29	32	29	32	32	31
fehlend	3	3	8	0	8	3	3	0	3	0	0	1
Mittelwert	2,000	2,483	1,500	2,656	2,542	3,345	2,241	2,813	2,414	3,031	1,625	3,226
Median	2,000	2,000	1,000	2,000	2,000	3,000	2,000	3,000	2,000	3,000	1,500	3,000
Standard-abweichung	1,0351	,9495	,5898	,8273	1,0206	1,1425	,9124	1,0298	1,0183	1,1496	,7071	1,1750

	Satz13.G – Satz13.C	Satz14.H – Satz14.C	Satz15.I – Satz15.C	Satz16.J – Satz16.C	Satz17.K – Satz17.C	Satz18.L – Satz18.C
Z	-2,619b	-3,906b	-2,829b	-3,144b	-2,560b	-4,606b
Asymp.Sig. (2-seitig)	,009	,000	,005	,002	,010	,000

¹ Die markierten Vergleiche verweisen auf signifikante Unterschiede.

	Satz19. D	Satz19. G	Satz20. D	Satz20. H	Satz21. D	Satz21. I	Satz22. D	Satz22. J	Satz23. D	Satz23. K	Satz24. D	Satz24. L
N gültig	32	32	32	32	32	28	32	30	31	31	31	31
fehlend	0	0	0	0	0	4	0	2	1	1	1	1
Mittelwert	2,531	3,375	2,469	3,031	2,750	1,786	2,594	3,367	2,613	3,290	2,419	3,000
Median	2,000	3,000	2,000	3,000	3,000	1,000	2,000	3,000	2,000	3,000	2,000	3,000
Standard- abweichung	1,0468	1,0999	,9499	1,0313	1,0776	1,0313	,9108	,9994	1,0223	1,0706	,8860	,9309

	Satz19.G - Satz19.D	Satz20.H - Satz20.D	Satz21.I - Satz21.D	Satz22.J - Satz22.D	Satz23.K - Satz23.D	Satz24.L - Satz24.D
Z	-3,681b	-3,350b	-3,281c	-4,104b	-3,285b	-3,368b
Asymp. Sig. (2-seitig)	,000	,001	,001	,000	,001	,001

	Satz25. E	Satz25. G	Satz26. E	Satz26. H	Satz27. E	Satz27. I	Satz28. E	Satz28. J	Satz29. E	Satz29. K	Satz30. E	Satz30. L
N gültig	28	30	31	32	31	32	30	32	31	32	31	31
fehlend	4	2	1	0	1	0	2	0	1	0	1	1
Mittelwert	2,964	3,533	3,000	3,156	2,161	3,188	1,400	2,813	2,226	3,000	2,645	3,323
Median	3,000	3,000	3,000	3,000	2,000	3,000	1,000	3,000	2,000	3,000	3,000	3,000
Standard- abweichung	1,1701	1,1059	1,0954	1,0809	,7788	1,0298	,7240	,9311	1,2030	1,0473	1,0181	1,0766

	Satz25.G - Satz25.E	Satz26.H - Satz26.E	Satz27.I - Satz27.E	Satz28.J - Satz28.E	Satz29.K - Satz29.E	Satz30.L - Satz30.E
Z	-3,116b	-,816b	-4,197b	-4,809b	-3,236b	-3,841b
Asymp. Sig. (2-seitig)	,002	,414	,000	,000	,001	,000

	Satz31. F	Satz31. G	Satz32. F	Satz32. H	Satz33. F	Satz33. I	Satz34. F	Satz34. J	Satz35. F	Satz35. K	Satz36. F	Satz36. L
N gültig	25	31	29	29	32	32	31	32	28	31	31	32
fehlend	7	1	3	3	0	0	1	0	4	1	1	0
Mittelwert	1,480	2,806	1,759	1,759	1,813	2,891	1,806	3,500	1,929	2,839	1,806	3,281
Median	1,000	3,000	2,000	1,000	1,500	3,000	2,000	4,000	2,000	3,000	2,000	3,000
Standard- abweichung	,7141	1,0139	,8305	,9876	1,0607	1,1621	,8725	1,2181	,8133	,9344	,8334	1,0846

	Satz31.G - Satz31.F	Satz32.H - Satz32.F	Satz33.I - Satz33.F	Satz34.J - Satz34.F	Satz35.K - Satz35.F	Satz36.L - Satz36.F
Z	-3,943b	-,030c	-3,680b	-4,620b	-3,793b	-4,356b
Asymp. Sig. (2-seitig)	,000	,976	,000	,000	,000	,000

Anhang 12: Kreuztabellen für den McNemar-Test

Anhang 12.A.: Kreuztabelle für die Fehlerangaben der Grammatik- und Aussprache-Fehlergruppen

			Fehlerangaben der Grammatik-Fehlergruppen		Gesamtsumme
			Fehler nicht erkannt	Fehler richtig erkannt	
Fehlerangaben der Aussprache-Fehlergruppen	Fehler nicht erkannt	Anzahl	177	585	762
		% in Fehlerangaben der Aussprache-Fehlergruppen	23,2%	76,8%	100,0%
		% in Fehlerangaben der Grammatik-Fehlergruppen	85,9%	62,2%	66,4%
	Fehler richtig erkannt	Anzahl	29	356	385
		% in Fehlerangaben der Aussprache-Fehlergruppen	7,5%	92,5%	100,0%
		% in Fehlerangaben der Grammatik-Fehlergruppen	14,1%	37,8%	33,6%
Gesamtsumme		Anzahl	206	941	1147
		% in Fehlerangaben der Aussprache-Fehlergruppen	18,0%	82,0%	100,0%
		% in Fehlerangaben der Grammatik-Fehlergruppen	100,0%	100,0%	100,0%

	Wert	Exakte Sig. (2-seitig)
McNemar-Test Anzahl der gültigen Fälle	1147	,000

Anhang 12.B.: Kreuztabelle für die Sicherheitsangaben der Grammatik- und Aussprache-Fehlergruppen

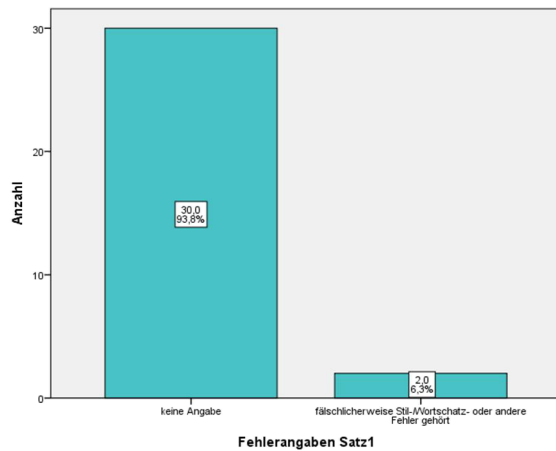
			Sicherheitsangaben Grammatik		Gesamtsumme
			,00	1,00	
Sicherheitsangaben Aussprache	,00	Anzahl	333	375	708
		% in Sicherheitsangaben Aussprache	47,0%	53,0%	100,0%
		% in Sicherheitsangaben Grammatik	81,6%	50,8%	61,8%
	1,00	Anzahl	75	363	438
		% in Sicherheitsangaben Aussprache	17,1%	82,9%	100,0%
		% in Sicherheitsangaben Grammatik	18,4%	49,2%	38,2%
Gesamtsumme		Anzahl	408	738	1146
		% in Sicherheitsangaben Aussprache	35,6%	64,4%	100,0%
		% in Sicherheitsangaben Grammatik	100,0%	100,0%	100,0%

Chi-Quadrat-Tests

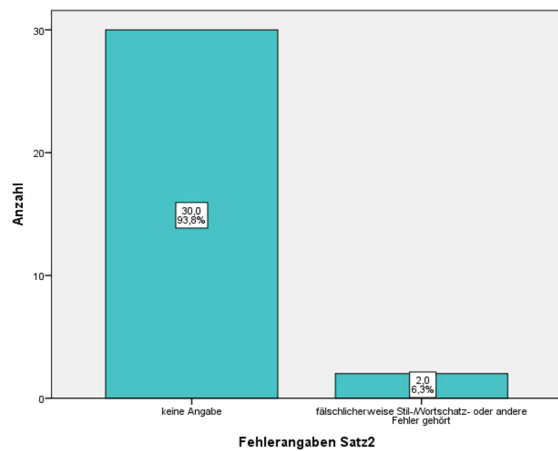
	Wert	Exakte Sig. (2-seitig)
McNemar-Test		,000
Anzahl der gültigen Fälle	1146	

Anhang 13: Fehlerangaben in den einzelnen korrekten Sätzen

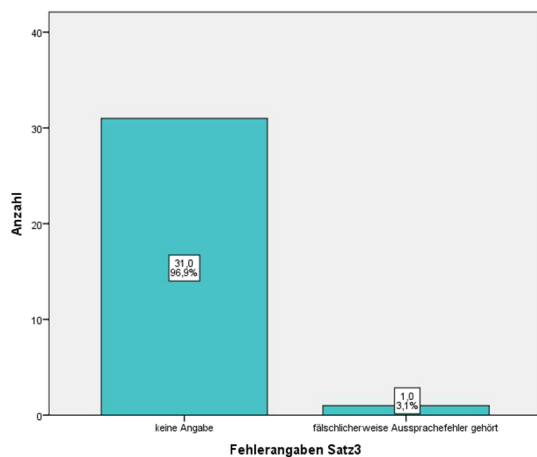
1. Ich habe leider keine Zeit.



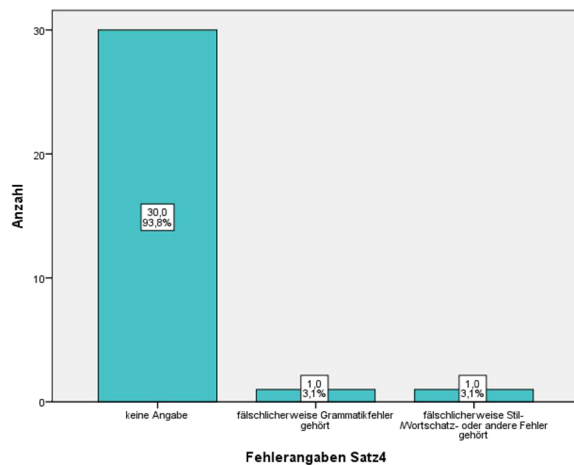
2. Das Justizlexikon steht im Regal.



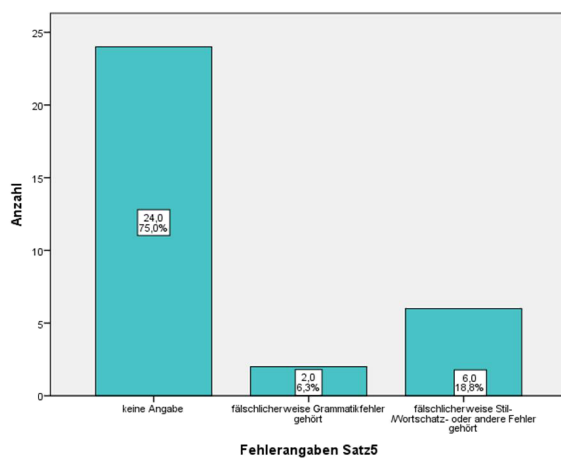
3. Er denkt nicht an seine Zukunft.
Sonne



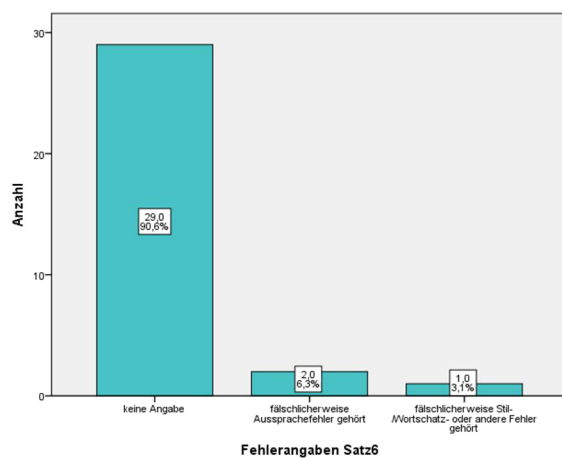
4. Meine Augen waren vom Glanz der
geblendet.



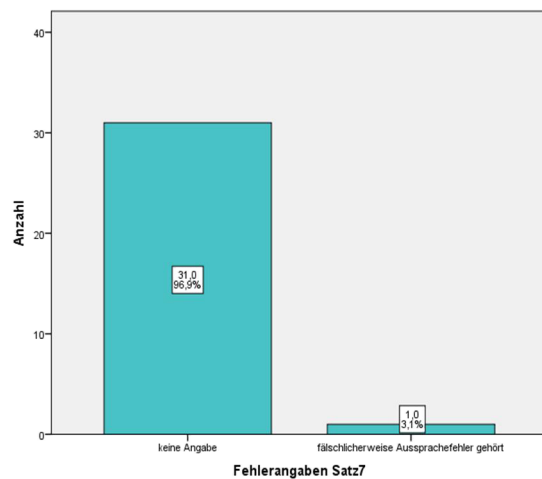
5. Neulich ist Benzin viel teurer geworden.



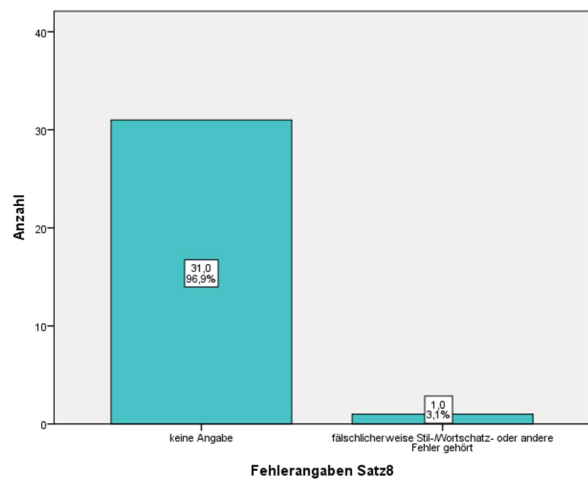
6. Alle wissen, dass die Touristen sich mit
Malaria infiziert haben.



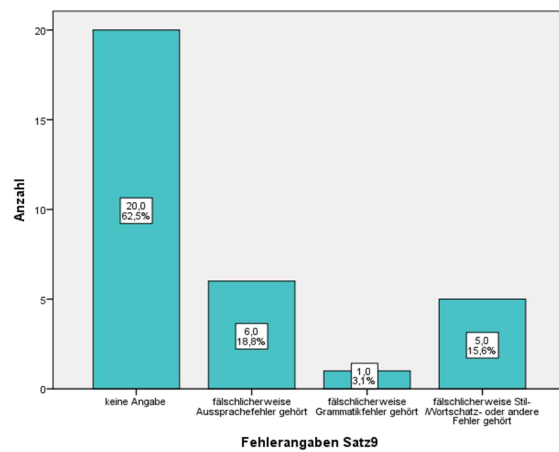
7. Meine Freundin ist naiv, aber nicht dumm.



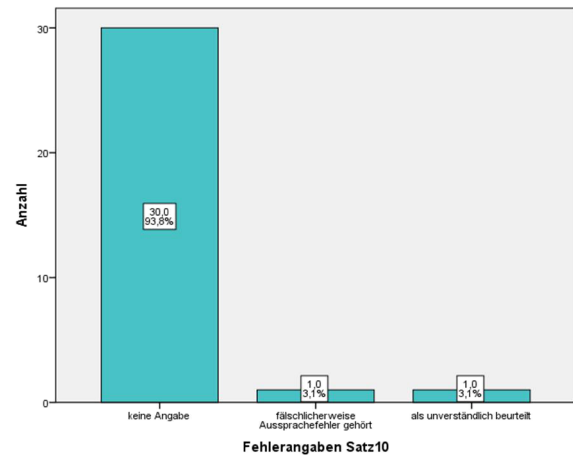
8. Der Wind blies mir ins Gesicht.



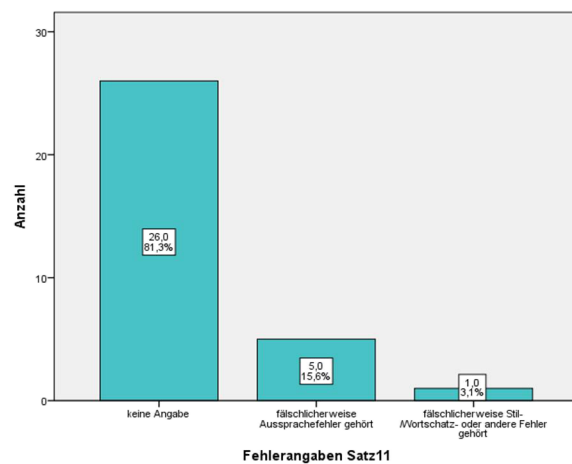
9. Er beteiligt sich immer aktiv am Unterricht.



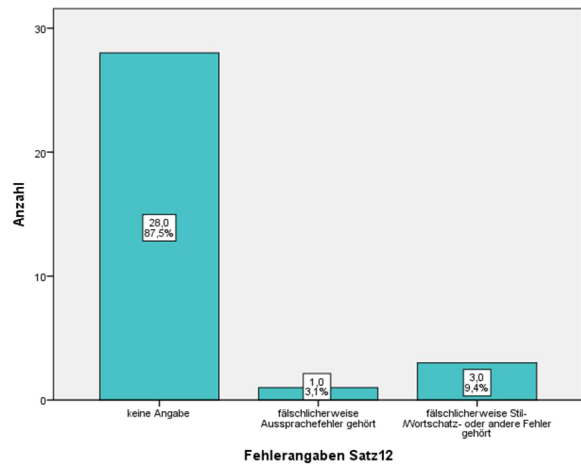
10. Einer der Hunde war sehr aggressiv und hat andere Hunde verletzt.



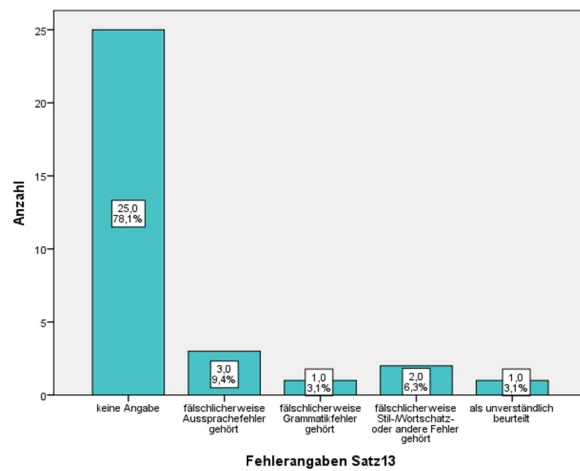
11. Leider wies die Rechnung einige Fehler auf.



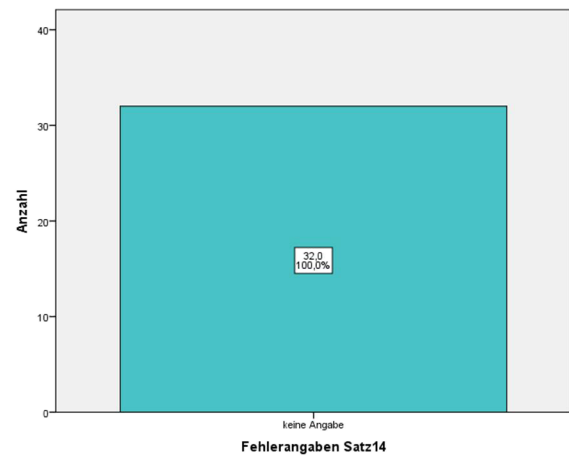
12. Wenn du dich mehr informieren willst, lies dieses Buch.



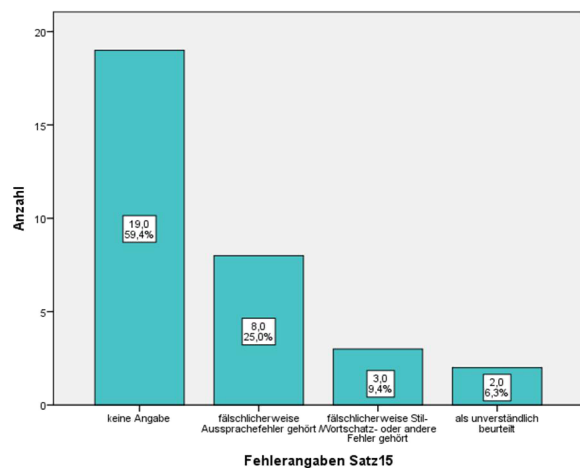
13. Wie lange kannst du deinen Atem anhalten?



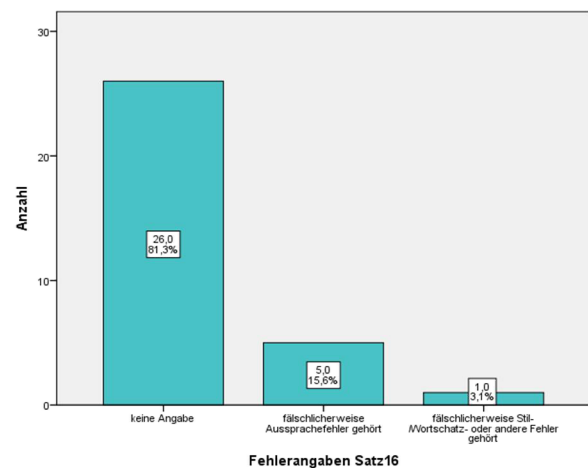
14. Sie ist die gute Seele in unserem Team.



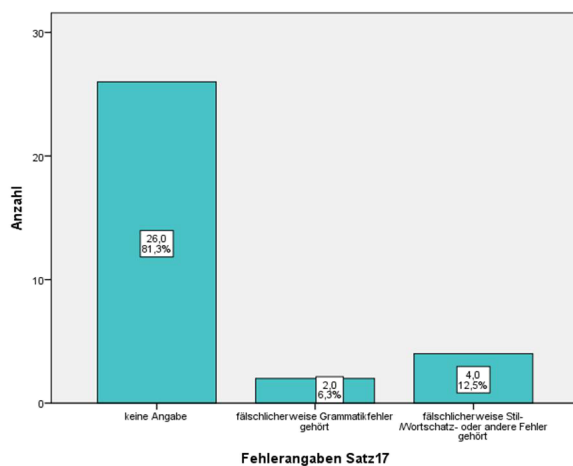
15. Mein Zahn ist sehr empfindlich und tut beim Kauen weh.



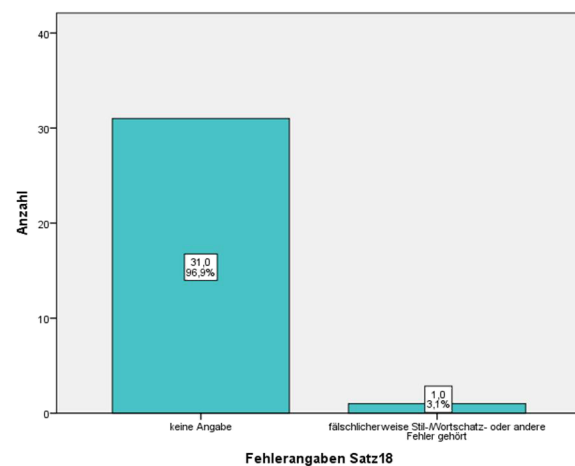
16. Es wird ewig dauern, bis er seine Schulden bezahlt hat.



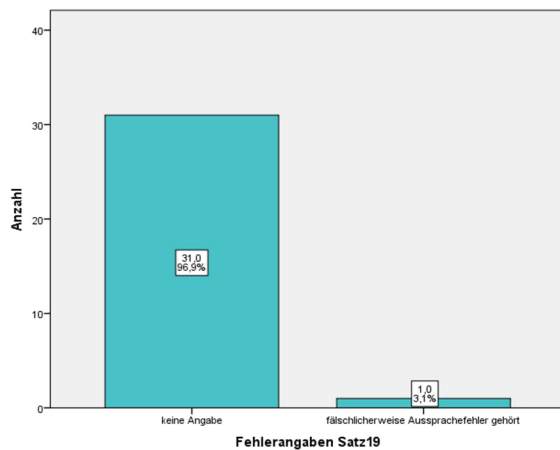
17. Am Wochenende sollte das Wetter schön werden.



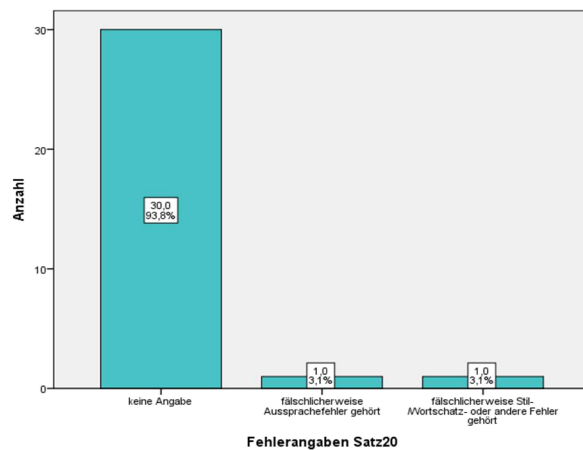
18. Die Kinder spielen mit dem Ball, obwohl sie es gar nicht dürfen.



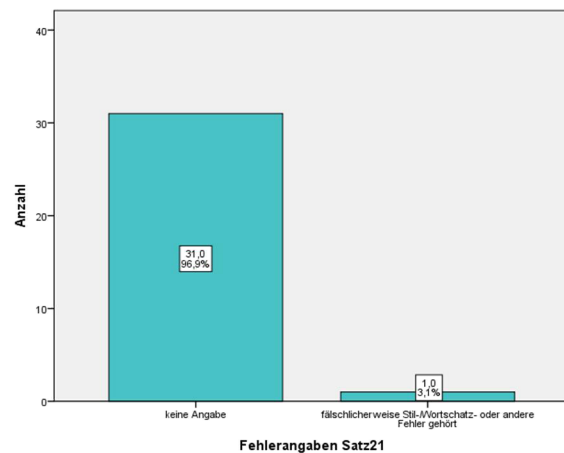
19. Die Gastgeber servierten ein üppiges Mahl.



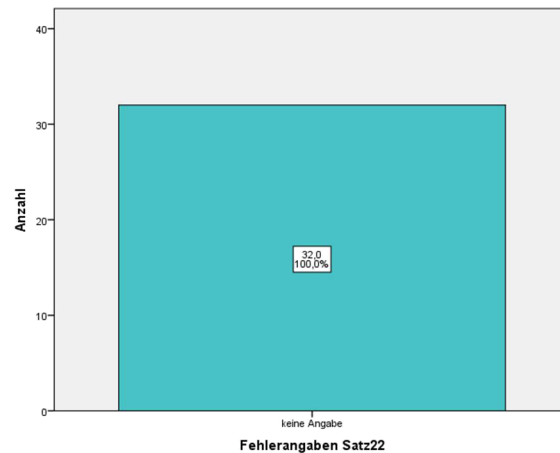
20. Kein Problem. Das nehme ich dir nicht übel.



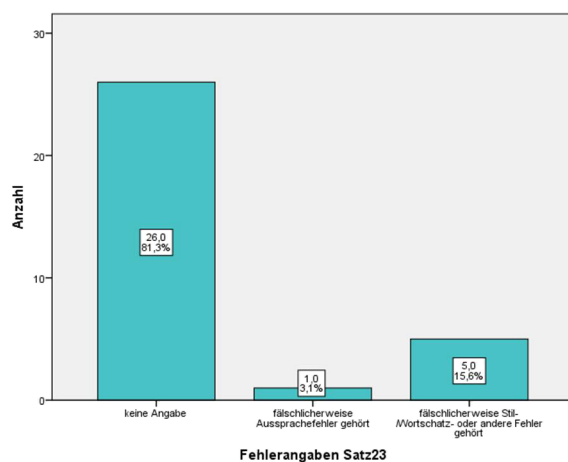
21. Sie sind sehr gütig zu mir.



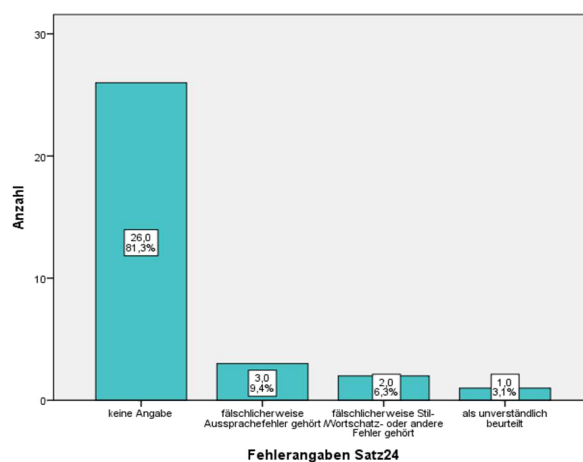
22. Sie hat sich gar nicht bemüht.



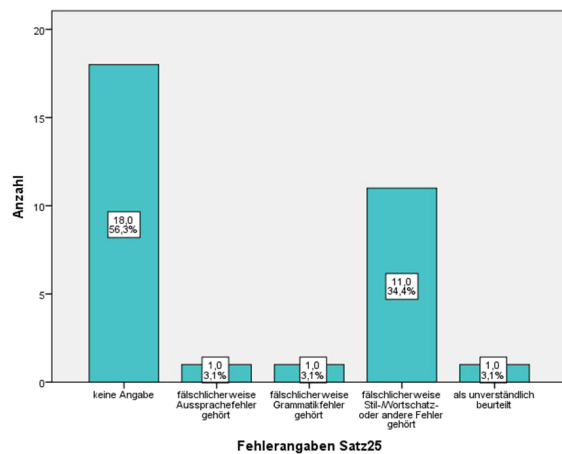
23. Jeden Tag treffen sich die Künstler.



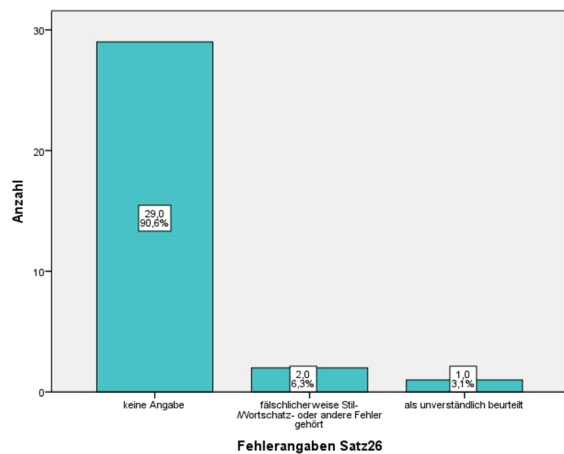
24. Auch wenn es hier sehr gemütlich ist, müssen wir bald wieder aufbrechen.



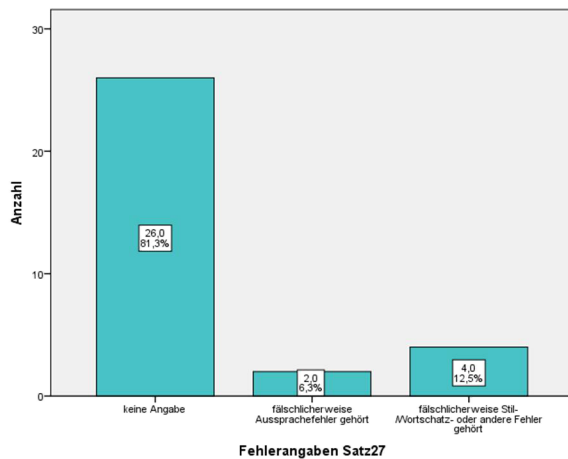
25. Deutsch ist für mich eine zweite Sprache.



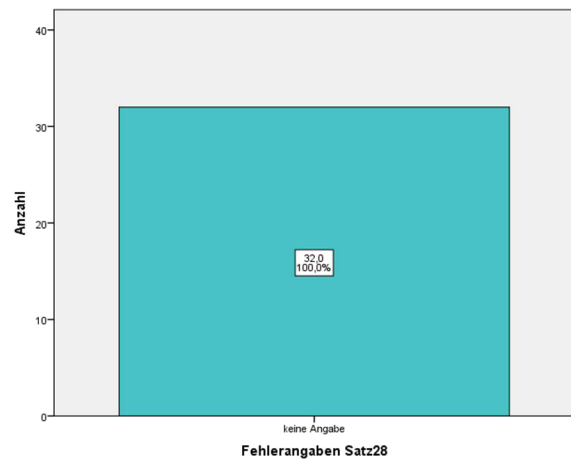
26. Wir sind auf der Straße.



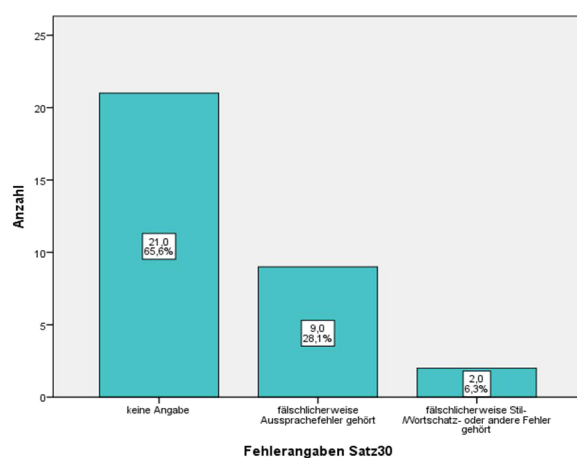
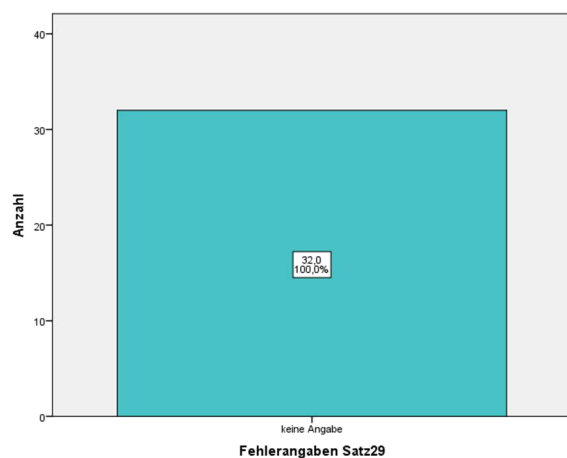
27. Ich habe Angst vor Schlangen.



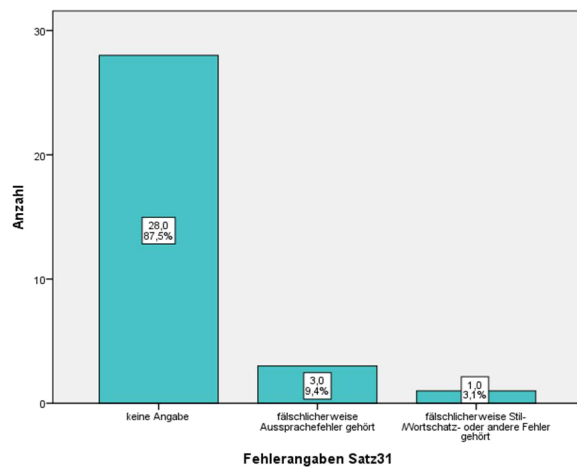
28. Die Verbrennungen haben ihr unerträgliche Schmerzen bereitet.



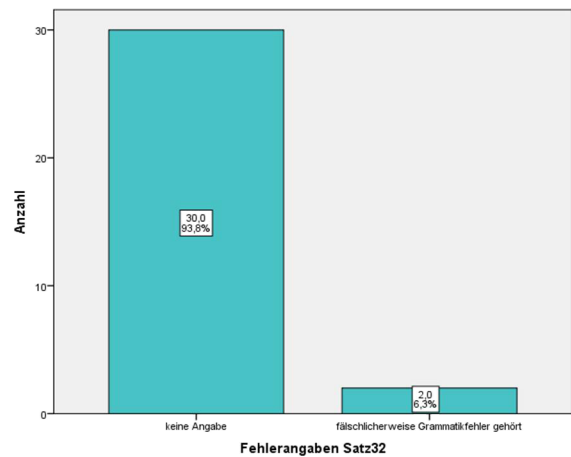
29. Letztes Jahr habe ich mein Studium beendet. 30. Obwohl er nicht so groß ist, kann er drei Meter hoch springen.



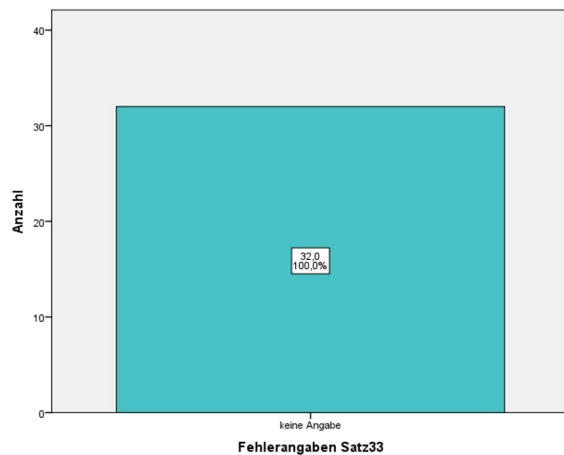
31. Wir treffen uns am Hauptbahnhof.



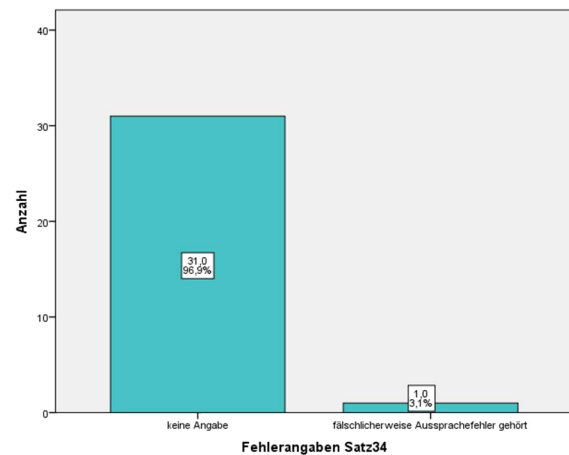
32. Die Kinder sind auf dem Spielplatz.



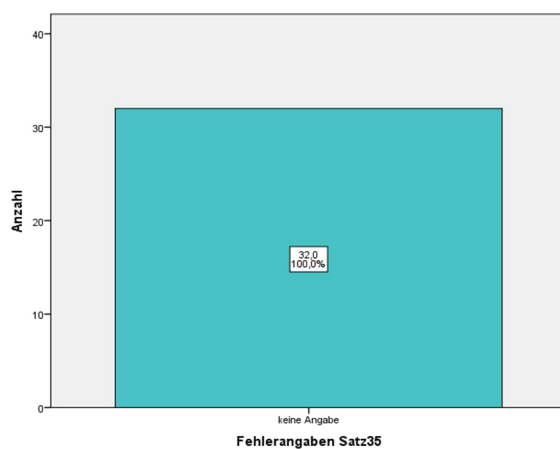
33. Meine Freundin ist immer noch im Krankenhaus.



34. Der Taxifahrer hat an der Bushaltestelle gehalten.



35. Als ausländischer Student muss man in Deutschland zuerst einen Sprachkurs besuchen.



36. Ich wusste nicht, dass das Studienkolleg hier in der Nähe ist.

